Statistical Analysis on Diffusion Tensor Estimation

Jiajia Yan BSc MSc

A thesis submitted in partial fulfilment of the
requirements of the University of Wolverhampton for
the degree of Doctor of Philosophy

January 2017

-
-

Signature……

Date……

Abstract

Diffusion tensor imaging (DTI) is a relatively new technology of magnetic resonance imaging, which enables us to observe the insight structure of the human body in *vivo* and non-invasively. It displays water molecule movement by a 3×3 diffusion tensor at each voxel. Tensor field processing, visualisation and tractography are all based on the diffusion tensors. The accuracy of estimating diffusion tensor is essential in DTI.

This research focuses on exploring the potential improvements at the tensor estimation of DTI. We analyse the noise arising in the measurement of diffusion signals. We present robust methods, least median squares (LMS) and least trimmed squares (LTS) regressions, with forward search algorithm that reduce or eliminate outliers to the desired level. An investigation of the criterion to detect outliers is provided in theory and practice. We compare the results with the generalised non-robust models in simulation studies and applicants and also validated various regressions in terms of FA, MD and orientations. We show that the robust methods can handle the data with up to 50% corruption. The robust regressions have better estimations than generalised models in the presence of outliers.

We also consider the multiple tensors problems. We review the recent techniques of multiple tensor problems. Then we provide a new model considering neighbours' information, the Bayesian single and double tensor models using neighbouring tensors as priors, which can identify the double tensors effectively. We design a framework to estimate the diffusion tensor field with detecting whether it is a single tensor model or multiple tensor model. An output of this framework is the Bayesian neighbour (BN) algorithm that improves the accuracy at the intersection of multiple fibres. We examine the dependence of the estimators on the FA and MD and angle between two principal diffusion orientations and the goodness of fit. The Bayesian models are applied to the real data with validation. We show that the double tensors model is more accurate on distinct fibre orientations, more anisotropic or similar mean diffusivity tensors.

The final contribution of this research is in covariance tensor estimation. We define the median covariance matrix in terms of Euclidean and various non-Euclidean metrics

taking its symmetric semi-positive definiteness into account. We compare with estimation methods, Euclidean, power Euclidean, square root Euclidean, log-Euclidean, Riemannian Euclidean and Procrustes median tensors. We provide an analysis of the different metric between different median covariance tensors. We also provide the weighting functions and define the weighted non-Euclidean covariance tensors. We finish with manifold-valued data applications that improve the illustration of DTI images in tensor field processing with defined non-weighted and weighted median tensors. The validation of non-Euclidean methods is studied in the tensor field processing. We show that the root square median estimator is preferable in general, which can effectively exclude outliers and clearly shows the important structures of the brain. The power Euclidean median estimator is recommended when producing FA map.

Symbols

b	diffusion weighting factor
$\boldsymbol{B}$	cholesky decomposition of tensor **D**
$\boldsymbol{D}$	diffusion tensor
d_A	power Euclidean distance
d_E	Euclidean distance
d_H	square root Euclidean distance
d_L	log-Euclidean distance
d_P	procrustes size-and-shape distance
d_R	Riemannian distance
g_i	diffusion encoding unit vector
K	Bayes factor
$\boldsymbol{M}$	net magnetisation vector
$\boldsymbol{S}$	diffusion signals
S_0	signal in absence of diffusion weighting
S_i	signal at the i^{th} diffusion gradient direction
$\boldsymbol{T}_A$	power Euclidean median covariacne tensor
$\boldsymbol{T}_E$	Euclidean mean covariance tensor
$\boldsymbol{T}_H$	square root median covariacne tensor
$\boldsymbol{T}_L$	log-Euclidean median covariacne tensor
$\boldsymbol{T}_P$	procrustes median covariacne tensor
$\boldsymbol{T}_R$	riemannian median covariacne tensor
v_1	eigenvector of tensor **D** corresponding to the largest eigenvalue
w	weighting function
ε_i	noise with independent identically distributed normal distribution $N(0, \sigma^2)$
σ	standard deviation of diffusion signals
φ	main fibre direction of tensor **D**

Acronyms/Abbreviations

ADC	apparent diffusion coefficient
AIC	Akaike information criterion
BIC	Bayesian information criterion
BN	Bayesian neighbours algorithm
CC	corpus callosum
CLLS	Cholesky linear least squares regression
CLMS	Cholesky least median squares regression
CNLS	Cholesky non-linear least squares regression
CT	computerised tomography
DSI	diffusion spectrum imaging
DTI	diffusion tensor imaging
DWI	diffusion weighted imaging
FA	fractional anisotropy
FS	forward search algorithm
GA	geodesic anisotropy
HARDI	high angular resolution diffusion weighted imaging
LLS	linear least squares regression
LMS	least median squares regression
LTS	least trimmed squares regression
MAD	median of all absolute distances
MD	mean diffusivity
MRI/NMRI	(nuclear) magnetic resonance imaging
NLS	non-linear least squares regression
ODF	orientation distribution function
PA	procrustes anisotropy
RA	relative anisotropy
RF	radio frequency
RMSE	root-mean-square error
ROI	region of interest
RSS	residuals sum of squares
VR	volume ration

Contents

List of Figures **vii**

List of Tables **xiii**

1 Introduction **1**

1.1 Magnetic Resonance Imaging 1

1.1.1 Theory 2

1.1.2 Diagnosis of clinical application of MRI 6

1.2 Diffusion tensor imaging 7

1.2.1 Diffusion tensor 8

1.2.2 Tensor estimations 10

1.2.3 Visualisation indices 13

1.2.4 Fibre tractography 14

1.2.5 Clinical applications 15

1.3 DTI tools 16

1.4 Thesis structure 25

2 Robust Algorithm For Tensor Estimation **27**

2.1 Introduction 27

2.2 Noise and artifacts 28

2.3 Robust statistics 29

2.3.1 Robust indices 30

2.3.2 Robust estimators 31
2.4 Methodologies 33
2.4.1 Least median squares (LMS) regression 33
2.4.2 Least trimmed squares (LTS) regression 33
2.4.3 Forward search (FS) algorithm for robust estimations 34
2.4.4 Test criterion 35
2.5 Simulations 37
2.6 Real data analysis 42
2.6.1 Robust estimation on a single tensor 42
2.6.2 Selection of outliers 51
2.6.3 DTI slices from three dimensions 55
2.6.4 Validation 64
2.7 Summary 66

3 Bayesian Model with Neighbourhood Priors on Multiple Fibres 69
3.1 Introduction 69
3.2 Multiple tensors models 72
3.2.1 High angular resolution diffusion weighted imaging 72
3.2.2 Mixture Gaussian 72
3.2.3 Diffusion spectrum imaging (DSI) 73
3.2.4 q-Ball Imaging 73
3.2.5 Spherical deconvolution methods 74
3.3 Methodology 75
3.3.1 Bayesian statistics 75
3.3.2 Single tensor model 77
3.3.3 Double tensor model 79
3.3.4 Occam's razor and model selection 83
3.4 Algorithm 85

3.4.1 Neighbours and voxels . . . 86
3.4.2 Starting point and phase portrait . . . 87
3.4.3 Bayesian neighbour's algorithm . . . 88
3.5 Simulation studies . . . 90
3.5.1 Dependence of angle between two tensors . . . 90
3.5.2 Dependence of mean diffusivity . . . 93
3.5.3 Dependence of fractional anisotropy . . . 94
3.5.4 Goodness of fit . . . 96
3.6 Real data analysis . . . 98
3.7 Summary . . . 98

4 Non-Euclidean Median Tensors Estimation 107
4.1 Introduction . . . 107
4.1.1 Tensors operations . . . 108
4.1.2 Euclidean and non-Euclidean metrics . . . 109
4.1.3 Fréchet Mean and mean covariance tensors . . . 111
4.2 Non-Euclidean estimators of median covariance tensors . . . 112
4.2.1 Estimators with reparameterisation . . . 113
4.2.2 Logarithm-based median estimators . . . 114
4.2.3 Procrustes estimator . . . 115
4.3 Weights . . . 116
4.3.1 Weighting functions . . . 116
4.3.2 Weighted metrics and median covariance tensors . . . 116
4.4 Algorithm . . . 117
4.5 Simulation studies . . . 119
4.5.1 Single tensor estimators with weights . . . 126
4.5.2 Smoothing . . . 127
4.5.3 Anisotropy study . . . 130

4.5.4 Smoothing with weighted median tensors 131
4.5.5 Interpolation with equal weights 132
4.6 Real data analysis 132
4.6.1 Smoothing with non-weighted median tensors 132
4.6.2 Smoothing with weighted median tensors 138
4.6.3 Interpolation 145
4.6.4 Validation 146
4.7 Summary 159

5 Conclusion and future work **161**
5.1 Conclusion 161
5.2 Future work 164
5.2.1 Tensor field segmentation and regularisation 164
5.2.2 Weights modification 164
5.2.3 Smoothing 165
5.2.4 Validation 165
5.2.5 Fibre tractography 165

Bibliography **167**

List of Figures

1.1 The hydrogen atom. 2

1.2 MRI machine components. 3

1.3 The Larmor precession. 3

1.4 The atoms without and with magnetic field. 4

1.5 DTI images in axial, sagittal and coronal view. 5

1.6 Net magnetic vector under applied field B_0. 6

1.7 An diffusion tensor ellipsoid with eigenvectors v_1, v_2 and v_3 of corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$. 9

2.1 The tensors D_1,D_2, D_3, D_4 and D_6. 38

2.2 The Root-mean-square error of FA, determinant and angle of D_1. 39

2.3 The Root-mean-square error of FA, determinant and angle of D_2. 40

2.4 The Root-mean-square error of FA, determinant and angle of D_3. 40

2.5 The Root-mean-square error of FA, determinant and angle of D_4. 41

2.6 The Root-mean-square error of FA, determinant and angle of D_5. 41

2.7 The Root-mean-square error of FA, determinant and angle of D_6 and its rotations with angles 15°, 30°, 45°, 60°, 75° and 90°. 46

2.8 A real DTI FA map estimated by LLS (top left), CNLS (top right), CLMS (bottom left) and LTS (bottom right) methods. 47

2.9 The tensor field of DTI image in Figure 2.8. 48

2.10 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (a) of Figure 2.8. 49

2.11 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (b) of Figure 2.8. 49
2.12 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (c) of Figure 2.8. 50
2.13 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (d) of Figure 2.8. 50
2.14 Percentages of voxels with outliers in each slice of DTI images from coronal, sagittal and transverse views. 51
2.15 Box plot of number of voxels with outliers in each slice of DTI images from coronal, sagittal and transverse views. 52
2.16 The number of outliers in each voxel of all voxels (left) and corrupted voxels (right). 52
2.17 The box plot and histogram of $\frac{z_i}{32}$. 53
2.18 The box plot and histogram of $\bar{z}_i$. 53
2.19 The signals under magnetics of voxel A, B and C. 54
2.20 The box plot and histogram of number of outliers in each voxel in set Ω. . . 54
2.21 The box plot and histogram of $\bar{z}_i$ in set Ω. 55
2.22 A DTI image in coronal view. 56
2.23 Number of outliers at each voxel of Figure 2.22. 56
2.24 The difference between each two FA maps. 57
2.25 A DTI image in sagittal view. 58
2.26 Number of outliers at each voxel of Figure 2.25. 58
2.27 The difference between each two FA maps. 60
2.28 A DTI image in transverse view. 61
2.29 Number of outliers at each voxel of Figure 2.28. 61
2.30 The difference between each two FA maps. 63
2.31 The corpus callosum region and its surrounding of a human brain. 64

3.1 Two fibres crossing (right), kissing (middle) and merging (left) of fibre bundles 70

3.2 The same tensors crossing at 15°, 30°, 45°, 60°, 75° and 90°, and their estimators in single tensor model 71

3.3 The neighbours of a voxel in 2D and 3D. 86

3.4 The three categories of voxels in 3D and its interior slice 87

3.5 The diagram of BN algorithm framework. 89

3.6 The simulated tensors D_1, D_{2a}, D_{2b} and D_{2c} in Section 3.5.1. 92

3.7 The estimated tensors field in Section 3.5.1. 92

3.8 The simulated tensors D_1, D_{2a}, D_{2b} and D_{2c}. 95

3.9 A slice of DTI images of a human brain in coronal view and its enlarge section. 99

3.10 The FA and v_1 map of estimated tensors using BN algorithm. The model selection is performed at each voxel. 100

3.11 The scatter plots of FA, MD and orientations between simulated tensors and estimated tensors between two diffusion tensors crossing at 30°, 60° and 90°. 102

3.12 The scatter plots of FA, MD and angle between simulated tensors and estimated tensors if a tensor has the MD size of 10, 3 times and the same size of another tensor. 104

3.13 The scatter plots of FA, MD and angle between simulated tensors and estimated tensors if one tensor is a linear, planar or spherical tensor. 106

4.1 Simulated metric space with variance 0.05(left), 0.1(middle) and 0.2(right). 120

4.2 Simulated 10×10 tensors and its FA and v1 map. 128

4.3 FA and v1 maps of median tensors using Euclidean(top left), power Euclidean(top right), log-Euclidean(middle left), root Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods. 129

4.4 The determinant of absolute values of tensors on each row. 130

4.5 Simulated 10×10 tensor field and its FA and v1 map. 131

4.6 FA and v_1 maps of median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods with inverse distances weights. 133

4.7 FA and v_1 maps of median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods with exponential weights. 134

4.8 Expanded 10×10 tensor field and its FA map. 135

4.9 Interpolation of 10×10 tensors FA map by log-Euclidean median tensors. . 136

4.10 FA maps of interpolation with median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods. 137

4.11 Real DTI images (axial view) and its enlarged splenium of corpus callosum region. 139

4.12 FA and v_1 maps of median tensors using Euclidean (top left), power Euclidean (top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian (bottom left) and discrete Procrustes (bottom right) distance methods. 140

4.13 A real DTI images (sagittal view) and its enlarged corpus callosum region. . 141

4.14 FA and v_1 maps of median tensors using Euclidean (top left), power Euclidean (top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian (bottom left) and discrete Procrustes (bottom right) distance methods. 142

4.15 A real DTI images (coronal view) and its enlarged corpus callosum region. . 143

4.16 FA and v_1 maps of median tensors using Euclidean (top left), power Euclidean (top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian (bottom left) and discrete Procrustes (bottom right) distance methods. 144

4.17 FA and v_1 maps of human brain (left) and the histogram of its FA values (right). 146

4.18 The tensor field of a human brain. 147

4.19 FA and v_1 maps of smoothing with inverse distance median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods. 148

4.20 Histogram of FA of smoothing with inverse distance median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods. 149

4.21 FA and v_1 maps of smoothing with exponential weighted median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods. 150

4.22 Histogram of FA of smoothing with exponential weighted median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods. 151

4.23 Smoothed tensors from Euclidean (top) and square root (bottom) median tensors methods. 152

4.24 FA and v_1 map of the human brain. 153

4.25 FA and v_1 map of smoothing with inverse distance median tensors using Euclidean (top) and log-Euclidean (bottom) methods. 153

4.26 FA and v_1 map human brain (top) and its smoothed image by Riemannian median tensors method (bottom). 154

4.27 FA and v_1 map of interpolation with median tensors using root Euclidean distance methods. 155

4.28 FA and v_1 maps of interpolation with median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods. 156

4.29 The selected voxels of corpus callosum. 157

4.30 The corpus callosum region and its surrounding of a human brain. 158

List of Tables

1.1 Glossary of terms in diffusion tensor Gaussian model. 10

1.2 The cost, main functions, implemented language, operation and system requirements, input data format, main users and comments of software tools that can be used for diffusion tensor imaging. 16

2.1 Fractional anisotropic, determinant and principal directions of D_1,D_2, D_3, D_4 and D_6. 38

2.2 The mean (median) of absolute value of difference in fractional anisotropy, mean diffusivity and angle of principal fibres among LLS, NLS, LMS and LTS methods on coronal view 69^{th} slice. 56

2.3 The mean (median) of absolute value of difference in fractional anisotropy, mean diffusivity and angle of principal fibres among LLS, NLS, LMS and LTS methods on sagittal view 65^{th} slice. 59

2.4 The mean (median) of absolute value of difference in fractional anisotropy, mean diffusivity and angle of principal fibres among LLS, NLS, LMS and LTS methods on transverse view 31^{th} slice. 62

2.5 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 2,000$. 65

2.6 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 4,000$. 65

2.7 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 5,000$. 65

2.8 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 1,0000$. 65

3.1 Jeffrey's scale of evidence for Bayes factors. 84

3.2 Kass and Raffery's scale of evidence for Bayes factors. 85

3.3 Summaries of three categories of voxels and the number of theirs single neighbours. 87

3.4 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between estimated tensors and their neighbours if they are crossing at 30°. 93

3.5 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between estimated tensors and their neighbours if they are crossing at 60°. 93

3.6 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and their neighbours if the two distinct fibres are crossing at 90°. 93

3.7 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors. 97

3.8 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors. 97

3.9 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors. 97

4.1 Summaries of metrics and mean covariance tensor estimators. 112

4.2 Summaries of Euclidean and non-Euclidean estimators, weighted metrics and weighted tensor estimators. 117

4.3 Measures of tensors properties with $\sigma = 0.05$. 120

4.4 Measures of tensors properties with $\sigma = 0.1$. 121

4.5 Measures of tensors properties with $\sigma = 0.2$. 122

4.6 Measures of efficiency with different distances ($\sigma = 0.05$). 123

4.7 Measures of efficiency with different distances ($\sigma = 0.1$). 124

4.8 Measures of efficiency with different distances ($\sigma = 0.2$). 125

4.9 Measures of tensors properties with $\sigma = 0.1$. 127

4.10 Root mean square error of fractional anisotropy, determinant of tensor and angle of principal directions of Euclidean mean, power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes median tensors tensors. 157

4.11 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated by Euclidean mean, power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes median tensors from raw data and noisy data with noise standard deviation $\sigma = 5,000$. 159

Chapter 1

Introduction

The research presented in this thesis focuses on the statistical analysis of tensor estimation in diffusion tensor imaging (DTI). DTI is an advanced technology of magnetic resonance imaging (MRI). It captures the details of biological microstructure in *vivo* and non-invasively.

1.1 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) and nuclear magnetic resonance imaging (NMRI) are medical imaging technologies that describe the details of brain anatomy using individual picture elements (pixels) with different intensities (brightness) in radiology. In 1946 Felix Block [13] and Edward Mills Purcell [85] both reported discoveries of nuclear magnetic resonance (NMR) in bulk matter, which formulated the underlying principles of MRI, for which they shared the Nobel Prize in physics in 1952. In 1973, Paul Lauterbur published the theory behind MRI [59], revolutionising clinical medicine. The first clinical image was also published that year and the first cross-sectional image, (of a living mouse), in the next year[59, 60]. Due to this work, the development of MRI was made possible. Lauterbur was awarded the Nobel Prize along with Peter Mansfield in physiology or medicine in 2003 for the development of MRI. Nuclear magnetics imaging was later renamed magnetic resonance imaging as people found the term 'nuclear' offputting. There are various coefficients in MRI, the signals of MRI are related to more than ten variables. Small deviations in each coefficient will have large variations on the attenuations.

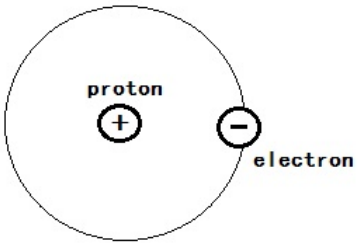


Figure 1.1 The hydrogen atom.

1.1.1 Theory

Nuclei with an odd number of protons can produce a small *magnetic moment* as a consequence of nuclear spin. It has direction and force of effect, thus it may be described by a vector. The size of a nuclear magnetic moment determines the strength of the MRI signal. Hydrogen has the simplest nucleus, having a single proton. It has the strongest magnetic moment. Figure 1.1 describes a hydrogen atom. As nearly two thirds of the human body is water, hydrogen nuclei is usually selected as the target in medical MRI. Each proton in the human body can be regarded as a small magnet. These small magnets are randomly orientated under normal circumstances.

There are four important components of MRI hardware: primary magnets, gradient magnets, radio frequency (RF) coils and computer systems. All MRI machine components are illustrated in Figure 1.2. The *primary magnetic field*, B_0, refers to the strength of the static permanent field. It is usually a constant $1.5-3.0$ Tesla in MRI. When the primary field is applied to the body, hydrogen atoms will align parallel or antiparallel (i.e. parallel but with opposite alignments) to the applied field. This is called *longitudinal magnetisation*. The protons spinning around the long axis of B_0, are termed *Larmor precession*. This is shown in Figure 1.3. The *Larmor frequency* is the frequency of Larmor precession given by:

$$w_0 = \gamma B_0 \tag{1.1}$$

where B_0 is the strength of the applied magnetic field and γ is a nuclei specific constant coefficient, called the gyromagnetic ratio. $\gamma = 42.6$ MHz/Tesla for hydrogen. The frequency changes in proportion to the magnetic field strength, at 1.5 Tesla, i.e. 63.9 MHz. when protons process together, it is called *in phase*. When protons process separately, they are said to be *out of phase*.

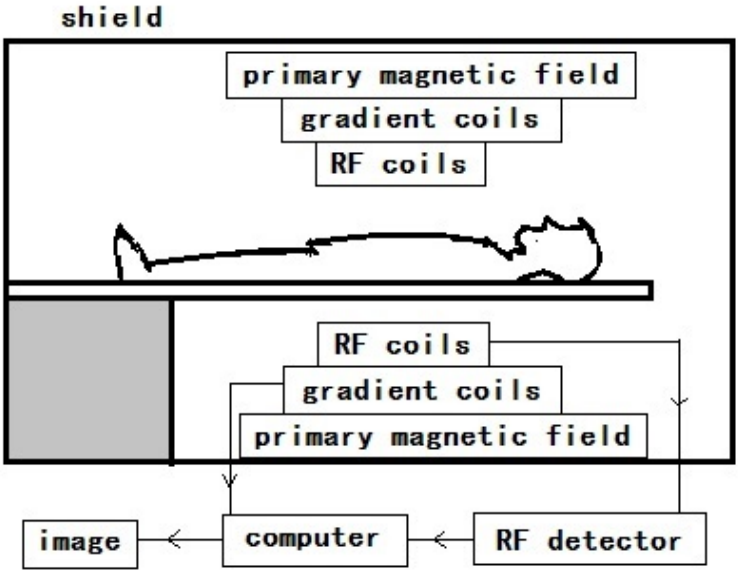


Figure 1.2 MRI machine components.

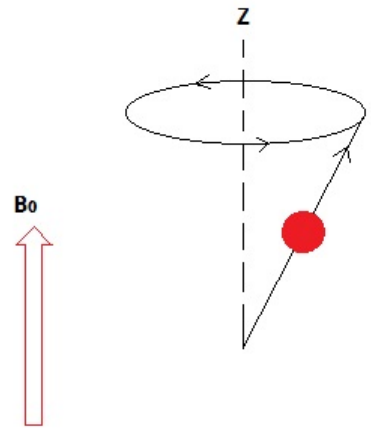


Figure 1.3 The Larmor precession.

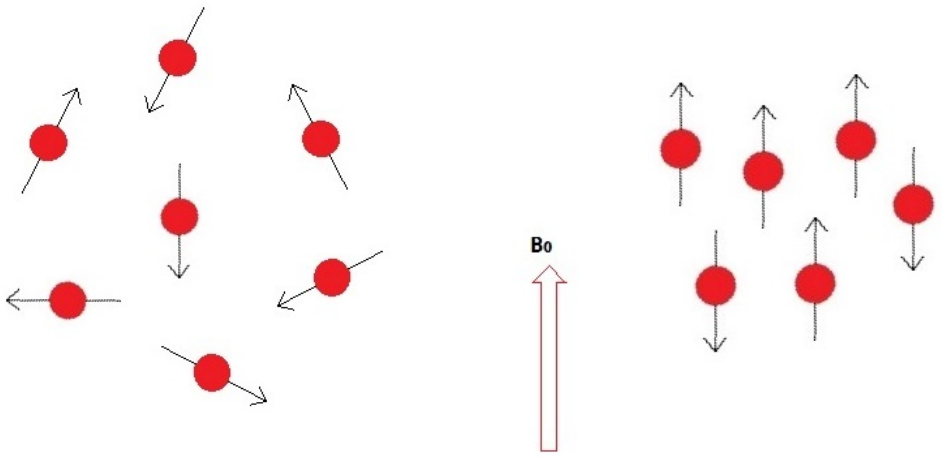


Figure 1.4 The atoms without and with magnetic field.

During the Larmor procession, there are two kinds of magnetic moment orientation: most protons are aligned along the direction of the magnetic field with low and stable potential energy; a small portion of protons are arranged antiparallel to the magnetic field with high potential energy. Figure 1.4 describes that the movement of atoms is randomly orientated under normal circumstances and become moving directionally under magnetic field. The remaining difference between these two is called *spin*. Spin generated by the residual magnetisation vector is called the *net magnetisation vector*, or *equilibrium macroscopic magnetic field oriented vector M*. Thus *M* is in the direction of the primary magnetic field. At a constant temperature, the ratio of the two elements depends on the intensity of the applied magnetic field.

Gradient coils generate a second magnetic field over the primary field. They are located within the bore of the primary magnet. There are three gradient coils, along with the X, Y and Z axis, and produce positive and negative pulsed fields. This allows spatial encoding for MRI images in X, Y and Z axis, i.e. *localisation*. In other words, it allows slice selection in these three directions by changing the procession frequency between slices. The X gradient produces an axial image, the Y gradient produces a sagittal image and the Z gradient produces a coronal image. Three DTI images from X, Y and Z gradients are shown in Figure 1.5.

The radio frequency or RF coils used to transmit the radio frequency pulse and receive signals. They are designed for specific body regions to give the best fit of head, body, knee,

Figure 1.5 DTI images in axial, sagittal and coronal view.

shoulder, wrist and ankle. An appropriate RF pulse is applied, which frequency is equal to the resonance frequency of hydrogen atom. There are two effects. First, some low energy parallel protons will flip to a high state decreasing longitudinal magnetisation. Second, the protons will process in phase gradually. In the MR coordinate system, the main magnetic field along the Z-axis direction is also known as the vertical axis, perpendicular to the main magnetic field direction of the XY plane. M spins around the Z axis with Larmor frequency. If applying an extra radio frequency (RF) pulse, perpendicular to B_0, M will deviate from the Z axis to the XY plane, and thereby form a *transverse magnetisation* vector. The Z axis deviation angle is called the *flip angle*. The flip angle is determined by the product of duration and strength of the RF pulse. The pulse which makes M transfer to the XY plane is called the *90 degree pulse*. The protons will also move in the same direction. When the external RF pulse stops, the transverse magnetisation vector M will gradually return to the Z axis. The *relaxation* is the time to reach an equilibrium state. It is measured in two directions:

- T_1, longitudinal – parallel to B_0 (Z axis)
- T_2, transverse – perpendicular to B_0 (XY plane)

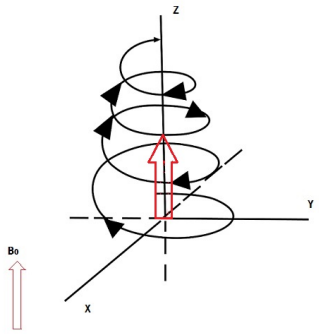


Figure 1.6 Net magnetic vector under applied field B_0.

The T_1 *relaxation* is also called *spin lattice relaxation*. It refers to the atoms return to low energy states after RF pulse. The protons that were in phase begin to decrease out of the Larmor frequency when RF pulse stops, is T_2 *relaxation*, or *spin spin relaxation*. The relaxation varies between different tissues and structures. The *net magnetic vector* is the sum of longitudinal and transverse magnetisation. It spirals around the Z axis with net precession. Changing the magnetic moment of the net magnetic vector, it releases energy in the form of RF signals, and restores to its original unaligned movement. The released three-dimensional coded RF signals received by the coil will be processed to 3D grey-scale MR images by the computer by performing an analog-to-digital conversion. Figure 1.6 describes the net magnetic vector under applied field B_0.

1.1.2 Diagnosis of clinical application of MRI

MRI is widely used in hospitals for medical diagnosis, especially for the diagnosis of sports-related injuries on soft tissue characteristics and blood vessel flow void effect [79]. In MRI, many brain regions may be clearly distinguished, such as thoracic and abdominal organs diseases, heart disease, cerebrovascular accidents and vascular diseases. It also can be used to track diagnosis and evaluation of obstacles and functional tumour. As there is no danger of radiation exposure, MRI is often used in the detection and diagnosis of the reproductive system, breast, pelvic and bladder diseases.

MRI is increasingly popular among clinicians and researchers. Nowadays, there are over 60 million cases of magnetic resonance imaging examination every year [35]. Compared to the computerised tomography (CT), MRI is a non-invasive, fast and accurate clinical

diagnosis. The MRI has an excellent resolution on soft tissues. Examinations of the bladder, rectum, uterus, vagina, bones, joints, muscles are better than CT [50].

There are two key elements in MRI: spatial resolution (pixel size) and contrast. The spatial resolution is usually 1 mm- 3 mm or even smaller in model MRI. Different areas of the brain are given different pixel intensities. A disadvantage of MRI is there is little contrast within the white matter in the imaging. The white matter includes axons connecting various regions of the brain. These axons tend to form bundles together with other axons. As the majority of the bundles have fluid-like homogeneous structure and similar chemical compositions, most of the bundles in the white matter cannot be identified by MRI. Only partial white matter regions are visible by conventional MRI, for example, the corpus callosum and the anterior commissure at the mid-sagittal level.

It has been reported that the number of incidents harming patient has increased [94]. Patients with cochlear implants and cardiac pacemakers, shrapnel and metallic foreign bodies in the orbits may be harmed by MRI examinations [24]. The significantly ferromagnetic objects are potential risks to persons and equipment within the magnet room. Interaction of the magnetic and radio frequency fields with such objects can lead to heating or torque of the object during an MRI. Loud noise, such as clicking, banging or beeping, is generated during operation of MRI by switching of field gradients. The sound pressure levels can reach 120 dB(A), which is equivalent to a jet engine at take-off. Therefore it may cause hearing damage on some patients [101]. Furthermore, there are doubts concerning the safety of MRI for women during their first three months of pregnancy.

With the evolution of technology, some specialised applications of MRI have been developed, for example, functional MRI, diffusion weighted imaging (DWI), diffusion tensor imaging (DTI), diffusion spectrum imaging (DSI), perfusion weighted imaging, magnetic resonance angiography, magnetic resonance spectroscopy, susceptibility weighted imaging and MR electrography. This research in based on statistical analysis of the diffusion tensor imaging. We will introduce DTI in the next section.

1.2 Diffusion tensor imaging

Diffusion Tensor Imaging (DTI) is a well-established technology of Magnetic Resonance Imaging (MRI). It displays an insight-structure of the biological tissue in the human body in vivo by directionally describing water molecular diffusion [15, 7].

In the mid-1980s, the underlying theory of diffusion MRI was developed. Le Bihan [62] reported the potential of diffusion MRI clinically. In 1990, it was reported by Michael Moseley [76] that water diffusion in white matter was anisotropic, i.e. the water diffuses in a preferable direction. The effect of diffusion on proton relaxation varies depending on the orientation of tracts relative to the diffusion gradient applied by the imaging scanner. Mosely first used the term 'Tensor' to describe the movement of the molecules [76]. This anisotropy effect became an attractive idea to map out the orientation in space of the white matter tracks. It assumes that the fastest diffusion direction indicates the major orientation of the fibres. Le Bihan et al. [62] used a two-dimensional approach to obtain colour maps of fibre orientation from the diffusion coefficients in X and Y directions. The theoretical analysis of DTI was developed by Basser [61] in 1994 and it has become an important imaging technique. MRI can be used to capture the local chemical and physical properties of water, i.e. its molecular diffusion and flow. DTI is an advanced MRI technique that directionally describes the water molecular diffusion using the diffusion tensor **D**. This diffusion process is thermal Brownian motion or random motion of water molecules. A diffusion that is even in every direction is termed *isotropic*. Water diffuses asymmetrically in the white matter. It tends to be directed along the fibre rather than crossing different tissues. This *anisotropic* property may be used to illustrate the principal fibre direction.

1.2.1 Diffusion tensor

A diffusion tensor, a covariance matrix related to the molecular displacement at a particular voxel in the brain, is in the non-Euclidean space of 3×3 positive semi-definite symmetric matrices. The diffusion tensor is estimated at each voxel in the brain, and is obtained by fitting a physically motivated model on measurements from the Fourier transform of the molecule displacements [3]. We define a diffusion tensor relating to the measured signals as

$$\begin{aligned} \ln\left(\frac{A(\mathbf{b})}{A(\mathbf{b}=\mathbf{0})}\right) &= -\sum_{i=1}^{3}\sum_{j=1}^{3} b_{ij}D_{ij} \\ &= -b_{xx}D_{xx} + 2b_{xy}D_{xy} + 2b_{xz}D_{xz} + b_{yy}D_{yy} + 2b_{yz}D_{yz} + b_{zz}D_{zz} \\ &= -Trace(\mathbf{bD}) \end{aligned} \tag{1.2}$$

where b_{ij} is a component of the symmetric b-matrix, **b**. The b-matrix is derived from DWI and summarises the attenuation effect of all gradient waveforms in all three directions x, y and z [7].

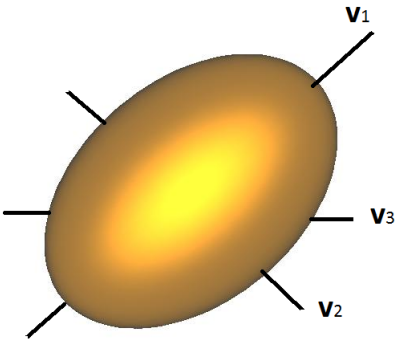


Figure 1.7 An diffusion tensor ellipsoid with eigenvectors v_1, v_2 and v_3 of corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$.

The DTI is a three-dimensional imaging, it requires a diffusion echo from at least six non-collinear, non-coplanar directions in order to provide enough information in Equation (1.2). Thus the diffusion tensor **D** is a symmetric matrix with six unknown parameters, the spectral decomposition of a diffusion tensor **D** is

$$D = U\Lambda U^T = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ v_3^T \end{pmatrix} \tag{1.3}$$

where **U** is 3×3 matrix whose i^{th} column is the eigenvector v_i of **D** and Λ is the diagonal matrix with diagonal elements λ_i, the corresponding eigenvalues. This eigensystem of **D** provides information to display the tissue microstructure in the human body. The symmetry property of **D** ensures that its eigenvalues are real, whereas positive semi-definiteness ensures that they are non-negative. In a particular voxel, the eigensystem of a diffusion tensor may be represented as an ellipsoidal surface with semi-major axis oriented along the v_1 direction and the semi-minor axes oriented along v_2 and v_3. The lengths of the axes in this diffusion ellipsoid are given by the square root of the corresponding eigenvalues of each eigenvector, with semi-major axis length $\sqrt{\lambda_1}$ and semi-minor axis lengths $\sqrt{\lambda_2}$ and $\sqrt{\lambda_3}$. The eigenvector v_1, corresponding to the greatest eigenvalue λ_1 characterises the principal direction of diffusivity. The determinant of **D** shows the volume of the tensor. Figure 1.7 shows a diffusion tensor ellipsoid.

The data at each voxel contain signals $\mathbf{S} = (S_0, S_1, \cdots, S_N)$ which are related to the Fourier transform of the displacements in axial directions $\mathbf{G}_j$ and the S_0 is the signal with no gradient.

Notation	Description
S_i	the signal amplitude in the i^{th} diffusion gradient direction
S_0	the signal in absence of diffusion weighting
b	diffusion-weighting factor
g_i	diffusion-encoding unit vector
ε_i	errors with i.i.d. $N(0,\sigma^2)$
$\mathbf{D}$	diffusion tensor

Table 1.1 Glossary of terms in diffusion tensor Gaussian model.

The Fourier transform in axial direction $\mathbf{g}$ of the multivariate Gaussian displacement is

$$F(\mathbf{g}) = \int i\sqrt{b}\mathbf{g}\mathbf{x} f(\mathbf{x}) d\mathbf{x} = \exp(-b\mathbf{g}^T\mathbf{D}\mathbf{g}) \tag{1.4}$$

For each voxel, the noise of the measured signal echo is denoted as ε_i, assuming independent and identically normally distributed with $N(0,\sigma^2)$. Thus the measured signal in a single voxel can be modeled as follows [7, 71, 8, 54]:

$$S_i = S_0 F(\mathbf{g}) = S_0 \exp(-b g_i^T D g_i) + \varepsilon_i \,, \qquad \text{where } i = 1, \cdots, N. \tag{1.5}$$

where S_i is the signal amplitude in the i^{th} diffusion gradient direction, S_0 is the signal in absence of diffusion weighting, b is the diffusion-weighting factor, g_i is the diffusion-encoding unit vector, $\mathbf{D}$ is a 3×3 positive semi-definite symmetric matrix. The diffusion tensor $\mathbf{D}$ describes the molecular displacement at a particular voxel in the brain [3, 15]. Table 1.1 summarises the meaning of each term in Equation (1.5).

Consequently, S_i are independent Gaussian variables with variance σ^2. The six coefficients of tensor $\mathbf{D}$ and the variance σ^2 of the noise are unknown parameters.

1.2.2 Tensor estimations

In a DT-MRI experiment, the measured signal in a single voxel can be modeled in Equation (1.5). The spectral decomposition of diffusion tensor $\mathbf{D}$ is [15][3]:

$$D = U\Lambda U^T = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ v_3^T \end{pmatrix} \tag{1.6}$$

where U is 3×3 matrix whose i^{th} column is the eigenvector v_i of $\mathbf{D}$ and Λ is the diagonal matrix whose diagonal elements λ_i are the corresponding eigenvalues. This eigensystem of $\mathbf{D}$ provides information to display the tissue microstructure of the human body. The eigenvector v_1, corresponding to the largest eigenvalue λ_1 characterises the principal direction of diffusivity. The determinant of $\mathbf{D}$ shows the volume of the tensor. We divide both sides of Equation (1.5) by S_0 and take the natural log of the equation and it becomes:

$$\log(S_i/S_0) = -bg_i^T Dg_i, \tag{1.7}$$

Thus it yields to a linearised problem, the components of the diffusion tensor are evaluated with the following objective function to be minimized:

$$f(\beta_{LLS}) = \frac{1}{2}\|y - X\beta_{LLS}\|^2 = \frac{1}{2}\sum_{i=1}^{N}\left(y_i - \sum_{j=1}^{6} X_{ij}\beta_j\right)^2 \tag{1.8}$$

where

$$y = [\log(S_i/S_0) \ \cdots \ \log(S_N/S_0)]^T \tag{1.9}$$

$$\beta_{LLS} = \begin{bmatrix} D_{xx} & D_{yy} & D_{zz} & D_{xy} & D_{yz} & D_{xz} \end{bmatrix}^T \tag{1.10}$$

$$X = -b\begin{pmatrix} \mathbf{g}_{1x}^2 & 2\mathbf{g}_{1x}\mathbf{g}_{1y} & 2\mathbf{g}_{1x}\mathbf{g}_{1z} & \mathbf{g}_{1y}^2 & 2\mathbf{g}_{1y}\mathbf{g}_{1z} & \mathbf{g}_{1z}^2 \\ \mathbf{g}_{2x}^2 & 2\mathbf{g}_{2x}\mathbf{g}_{2y} & 2\mathbf{g}_{2x}\mathbf{g}_{2z} & \mathbf{g}_{2y}^2 & 2\mathbf{g}_{2y}\mathbf{g}_{2z} & \mathbf{g}_{2z}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{g}_{Nx}^2 & 2\mathbf{g}_{Nx}\mathbf{g}_{Ny} & 2\mathbf{g}_{Nx}\mathbf{g}_{Nz} & \mathbf{g}_{Ny}^2 & 2\mathbf{g}_{Ny}\mathbf{g}_{Nz} & \mathbf{g}_{Nz}^2 \end{pmatrix} \tag{1.11}$$

where N is the number of diffusion gradient directions, β is a vector representation of the diffusion tensor, y is the $n \times 1$ vector of responses, X is an $n \times p$ full-rank matrix of known constants, with i^{th} row x_i^T, and β is a vector of p unknown parameters. The errors ε_i are independent and identically normally distributed with $N(0, \sigma^2)$.

The solution to the LLS problem may also be written explicitly as [54]:

$$\hat{\beta}_{LLS} = [X^T X]^{-1} X^T y = X^+ y \tag{1.12}$$

where $\hat{\beta}$ denotes the linear squares estimate of β and X^+ denoted the inverse of X. The LLS method is widely used, although the LLS estimates may not satisfy the positive semi-definiteness of tensors.

The non-linear least squares method can be also used to estimate the diffusion tensor. The minimising objective function expressed as follows [54, 55]:

$$f(\beta_{NLS}) = \sum_{i=1}^{N} \left[S_i - S_0 \exp\left(\sum_{j=1}^{N} X_{ij}\beta_j \right) \right]^2 \tag{1.13}$$

The diffusion tensor is a positive semi-definite symmetric matrix in DTI, i.e. The eigenvalues λ_i of Equation (1.6) are real and positive. The design matrix X is naturally assumed that the eigenvalues are real but regardless of the positive condition. The diffusion tensor parameter vector, β, can be reparameterised to add the positive semi-definite property on X. One such approach is the Cholesky decomposition [84]. The diffusion tensor can be expressed as the form

$$D = C^T C, \qquad C = \begin{pmatrix} c_1 & c_4 & c_6 \\ 0 & c_2 & c_5 \\ 0 & 0 & c_3 \end{pmatrix} \tag{1.14}$$

C is an upper triangular matrix with non-zero diagonal elements. Consequently, rewrite the C as a vector-valued function ρ and the diffusion tensor parameter vector β becomes

$$\rho = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 & c_5 & c_6 \end{bmatrix}^T;$$

$$\beta(\rho) = [c_1^2,\ c_2^2 + c_4^2,\ c_3^2 + c_5^2 + c_6^2,\ c_1 c_4,\ c_2 c_5 + c_4 c_6,\ c_1 c_6]^T \tag{1.15}$$

Introducing the Cholesky parameterisations in the Equation (1.13), the objective function for Cholesky non-linear least squares (CNLS) estimation is written as follows:

$$f(\rho_{CNLS}) = \sum_{i=1}^{N} \left[S_i - S_0 \exp\left(\sum_{j=1}^{N} X_{ij}\beta(\rho)_j \right) \right]^2 \tag{1.16}$$

Note that the CNLS estimate is equivalent to NLS estimate when the NLS estimate is positive semi-definite.

$$f_{CLLS}(\rho) == \frac{1}{2} \sum_{i=1}^{N} (y_i - \sum_{j=1}^{6} X_{ij}\beta(\rho)_j)^2; \tag{1.17}$$

1.2.3 Visualisation indices

There are some important visualisation indices used in DTI, which will be illustrated explicitly.

Mean Diffusivity (MD)

The Mean Diffusivity (MD), or trace, which measures the overall displacement of molecules, is the mean of eigenvalues. These measures are commonly used clinically to localise white matter lesions that do not show up on other forms of clinical MRI.

$$MD = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \tag{1.18}$$

Fractional Anisotropy (FA)

The Fractional Anisotropy (FA) describes the fraction of the 'magnitude' of **D** that can be ascribed to anisotropic diffusion [15]. FA is defined by

$$FA = \frac{\sqrt{3(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}}{\sqrt{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{1.19}$$

The value of FA indicates the shape of tensors. For example, a "cigar" shaped prolate ellipsoid has a strongly linear anisotropy, a "flying saucer" or oblate spheroid represents diffusion in a plane, and a sphere is indicative of isotropic diffusion, equal in all directions.

Relative Anisotropy (RA)

Relative Anisotropy (RA) is an alternative 'magnitude' measurement. It is a normalised standard deviation of the eigenvalues and is given by

$$RA = \frac{\sqrt{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}}{\sqrt{3MD}} \tag{1.20}$$

Volume Ratio (VR)

The volume ratio is calculated as follows

$$VR = \frac{\lambda_1 \lambda_2 \lambda_3}{MD^3} \tag{1.21}$$

1.2.4 Fibre tractography

The symbolic (or geometric) display methods are based on the use of different types of a glyph to display the spatial distribution of the anisotropy and the principal direction. Arrow, ellipsoid and other combined objects are the important indices.

$$c_l = \frac{\lambda_1 - \lambda_2}{\lambda_1} \tag{1.22}$$

$$c_p = \frac{\lambda_2 - \lambda_3}{\lambda_1} \tag{1.23}$$

$$c_s = \frac{\lambda_3}{\lambda_1} \tag{1.24}$$

$$c_l + c_p + c_s = 1 \tag{1.25}$$

The most attractive tool for visualising and analysing white matter fibre tracts is tractography. Fibre tractography was initiated by Mori [75], using the tracking fibres in principle direction of diffusion tensor. It is based on the line propagation technique . For a given start point, *seed*, P_0, the series of node coordinates P_i on the trajectory are determined as follows:

$$P_{i+1} = P_i + \theta d_i(D(P_i)) \tag{1.26}$$

where θ is a scalar value of the propagation step distance, d_i is the unit vector of propagation direction depending diffusion tensor $\mathbf{D}$ at the location P_i.

The imaging of this property is an extension of diffusion MRI. If a series of diffusion gradients (i.e. magnetic field variations in the MRI magnet) are applied that can determine at least 3 directional vectors (the use of 6 different gradients is the minimum and additional gradients improve the accuracy of "off-diagonal" information), it is possible to calculate, for each voxel, a tensor (i.e. a symmetric positive definite 3×3 matrix) that describes the 3-dimensional shape of diffusion. The fibre direction is indicated by the tensor's main eigenvector. This vector can be color-coded, yielding a cartography of the tracts' position and direction (red for left-right, blue for superior-inferior, and green for anterior-posterior). The

brightness is weighted by the fractional anisotropy which is a scalar measure of the degree of anisotropy in a given voxel.

1.2.5 Clinical applications

The principal application is in the imaging of white matter where the location, orientation, and anisotropy of the tracts can be measured.

White matter

Diffusion tensor imaging can be used to tract whole fibres within white matter (e.g. the corticospinal tract, through which the motor information transit from the motor cortex to the spinal cord and the peripheral nerves), especially fibre orientation and strength. By detecting deficit in white matter, tractography is a powerful tool, such as in ageing. It is reported that DTI has widespread potential in cognitive neuroscience and neurobiology.

Strokes

Acute brain injury is one of the important applications of DTI. In the acute phase of ischemia, the FA is increased and apparent diffusion coefficient (ADC) is decreased. In the chronic phase of ischemia, ADC is higher than normal and FA is lower in the infarcted area. The severity of strokes can be assessed and acute ischemic changes can be distinguished from chronic ischemic changes. This benefits the treatment of strokes.

Brain tumors

The gliomas and solitary metastasis in the brain parenchyma are clearly demonstrated by DTI with significantly higher mean diffusivity and lower FA values.

Demyelinating versus dysmyelinating disorders

Diffusion isotropy is present in demyelinating diseasea such as Krabbe disease or Alexander disease, while diffusion anisotropy is present in demyelinating disorders such as Pelizaeus-Merzbacher disease. This contrast is used to distinguish the demyelination from demyelination.

1.3 DTI tools

In this section, we will review the current DTI tools in Table 1.2. 3D slicer [81] is mainly used in hospitals. Brain Vovager QX [33] is required for partial licenses. The other softwares in DTI are free and academically, including: AFNI [23], AMIDE [68], BioImageSuite [77], BrainParser [67], CATNAP [58], CAMINO [22], DipY [32], DoDTI [78], DtiStudio [53], DSIStudtio [57], DTI-TK [109], DTI-Query and DTI-CINCH [96], DTI-Toolbox [34], ExploreDTI [64], FreeSurfer [28], FSL-FDT [99], HAMMER [95], ImageJ [74], IMPATIENT MRI [111], iBrain [98], JIST [69], MIPAV [11], MedINRIA [104], MrDiffusion [], Probtrackx [16], SepINRIA [100], SATURN [17], SPM and toolboxes [6], STAMPS [14], TBSS [99], TracTor [20], TrackVis [108] and TORTOISE [83].

Table 1.2 The cost, main functions, implemented language, operation and system requirements, input data format, main users and comments of software tools that can be used for diffusion tensor imaging.

No.	Software	Cost	Main functions	Implemented Language	Operation and System Requirements	Input data format	Main users	Comments
1	3D Slicer	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	C++	Windows, Linux and Mac OS,	BSD-style	Hospitals	User-friendly. Faster information exchange. Not a single software platform. Built on a set of powerful and widely used software components (Tcl/Tk, VTK, ITK)
2	AFNI	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	ANSI C	Unix X11 Motif systems, including SGI and Linux		Academic, all FMRI researchers at the Medical College of Wisconsin	Not specific to the stereotaxic transformation problem. Does not form a new 3D array with the transformed data from datasets.
3	AMIDE	Free under the terms of the GNU General Public License (GPL)	Quantitative analysis	C	Mac OS, Linux, and Windows.	XML (eXtensible Markup Language) based directory format	Academic	Provide variety of additional features. Simplified interface. Slice viewing is computationally expensive.

4	BioImage Suite	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis		Windows, Linux and Mac Os X	NIH funded	Academic	
5	BrainParser	Free	Quality control and preprocessing, Processing and visualisation	Java	Linux and Windows		Academic	Highly adaptive. Efficient computation. The variations are intrinsic. Numbers of training data are small.
6	BrainVo-yager QX	Partial licenses needed	Processing and visualisation, Quantitative analysis		Windows, Linux and Mac Os X		Academic	
7	CATNAP	Free	Quality control and preprocessing, coregistration, change, and tensor-tackling	Java	Matlab, Signal Processing Toolbox		Academic	Result of CATNAP is compatible with DTIStudio, FSL, and other tensor analysis packages. Accessible from other programming environments
8	CAMINO	Free	Quality control and preprocessing, Processing and visualisation,	Java	Unix X11 Motif systems, including SGI, Solaris, Linux, Windows,		Academic	

			Quantitative analysis		Matlab			
9	DipY	Free	Processing and visualisation	Python	Windows, Linux, Mac O S X		Academic	
10	DoDTI	Free	Quality control and preprocessing, Processing and visualisation,	Matlab	Windows, Linux, Mac and SUN Matlab Toolkit,	Accept a mesh data format	Academic	DVTK, .grad
11	DTIStudio	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	C++, OpenGL	Windows		Academic	
12	DSIStudio	Free	Quality control and preprocessing, Processing and visualisation	C++, Matlab	Windows	Matlab	Academic	
13	DTI-TK	Free	Quality control and preprocessing, Processing and visualisation	C++	Mac	Toolkit NIfTI	Academic	Support a standard-based IO file

14	DTI-Query and DTI-CINCH	Free	Processing and visualisation	Query language	Mac, Linux, Cygwin, Windows, AFNI software package		Academic	
15	DTI-Toolbox	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	IDL	Windows, OS X, Linux, Solaris		Academic	Only can display up to 2D image
16	ExploreDTI	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	Matlab	Windows, Linux and Mac Os X	Dicom, Analyze, NIFTI and Matlab Toolkit	Academic	
17	Free-surfer	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis		Linux and Mac Os X	Virtualbox needed, .dmg file	Academic	

18	FSL-FDT	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	R	Windows, Mac		Academic	
19	HAMMER	Free	Quality control and preprocessing	C	Linux and Windows.		Academic	
20	ImageJ	Free	Quantitative analysis	Java	Linux, Mac OS, and Windows.		Academic	
21	IMPATIENT MRI	Free	Quality control and preprocessing (GPU-based advanced image reconstruction)	C++	Linux, Mac OS, and Windows.		Academic	
22	iBrain	Free	Processing and visualisation (tensor estimation and tractography)		Windows, Linux and Mac Os X	Matlab or SPM toolbox	Academic	
23	JIST	Free	Quality control and preprocessing,	Java	Windows, Linux and Mac Os X	Java-based Toolkit GUI	Academic	

			Processing and visualisation		Sun Java 1.6 and above			
24	MIP A V	Free	Processing and visualisation, Quantitative analysis	Java	Mac OS, Linux, and Windows.		Academic	Supports all major medical image formats, and provides many visualisation and data manipulation tools for 2D and 3D images or image series
25	MedINRIA	Free	Processing and visualisation, Quantitative analysis	C++	Windows, MacOS and Linux		Academic	Construct 2D to 4D image visualisation, diffusion image processing, segmentation of images, filtering of images and registration of images.
26	MrDiffusion	Free	Processing and visualisation, tensor estimation, tractography		Windows, Linux and Mac Os X	Matlab and VISTASOFT need	Academic	
27	Probtrackx	Free	Processing and visualisation,	R	Windows, Mac FSL.		Academic	Probabilistic tracking implemented as a part of FDT
28	SepINRIA	Free	Processing and visualisation, Quantitative analysis	C#	Windows XP, Windows Vista, Linux Fedora Core , Mac OS X		Academic	Multiple sclerosis. Visualisation of 2D and 3D images, comparison of different sequences alignments, quantitative lesion burden evaluation, comparison between a segmentation and a segmentation of reference and

								evaluation of the brain atrophy.
29	SATURN	Free	Processing and visualisation, Quantitative analysis, tensor estimation	C++	Linux, Mac OS, and Windows	DICOM, Kretz, ni	Academic	Neuroimaging and incorporates a complete arrangement of representation capacities to scan and investigate proficiently DTI information and one of the capable devices for finding reason.
30	SPM and toolboxes	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	C	MATLAB		Academic	Concocted for measurable parametric mapping of PET and fMRI information. It is one of various bundles that give a comparable system, and in addition some fresher usefulness that separates from this.
31	STAMPS	Free	Quantitative analysis	tcl or tk	Linux, Mac OS, and Windows.	Multiple MR image types	Academic	Enables the automation of nonlinear registration for a large image set and for multiple MR image types. It also contains novel MRI post-processing features
32	TBSS	Free	Quantitative analysis	R	Windows, Mac, Matlab R2009a, MRIcro and DTIStudio vuTools, VMware virtual		Academic	Improve the sensitivity, objectivity and interpretability of the analysis Solve the alignments and smoothing issue arise from FA. Automate investigation on whole brain without specification of interested tracts.

					machine, VMware Player and Winzip			
33	TracTor	Free	Processing and visualisation	R	Unix		Academic	Reading, writing and manipulating MR images
34	TrackVis	Free	Processing and visualisation, Quantitative analysis	C++,using Qt 4.2 for the GUI and Visualisation Toolkit 5.0	Mac OS, Linux, and Windows.	Diffusion tensor data , high angular resolution diffusion imaging (HARDI) data as well as diffusion spectrum imaging (DSI) data and Q-Ball imaging data.	Academic	Reconstruction, fiber tracking, analysis and visualisation
35	TORTOISE	Free	Quality control and preprocessing, Processing and visualisation, Quantitative analysis	C++	Linux and Mac platforms.		Academic	Composed of two modules named DIFF PREP and DIFF CALC. DIFF PREP

1.4 Thesis structure

This thesis is divided into five chapters. In Chapter 1, we review the background of MRI and DTI. The diffusion tensor Gaussian model and tensor eigensystem are illustrated for later chapters. The traditional estimation, linear and non-linear least squares methods are introduced. Visualisation indices, mean diffusivity, fractional anisotropy, relative anisotropy and volume ratio are described. Fibre tractography and clinical applications are also illustrated in the chapter. The various commercial and non-commercial software for DTI are summarised with their main functions, implemented language, input data format and operation and system requirements.

In Chapter 2, we review possible noises and artifacts that may result in outliers in estimation. The robust theory is briefly introduced and some robust indices and estimators are described. We proposed robust regressions with forward search algorithm to reduce or eliminate outliers to the desired level. The results are compared with estimators from linear least squares and non-linear least squares methods in simulation studies and applicants. Simulation studies are carried in three cases: different fractional anisotropy, volume and orientation of the tensors to compare the root-mean-square errors of the properties of the tensors. A series of outliers are added in the simulation to test the performance of the algorithm. In the real data, we first compare the estimators of a single diffusion tensor and read the pattern of its neighbouring tensors. Then we analyse the real data arising in the diffusion signals and conclude the appropriate outliers selection methods and threshold values. The comparisons are applied to three DTI images from the coronal, sagittal and transverse view resumptively. The absolute value of difference in terms of FA, MD and orientation between the various estimators are shown in tables and figures. We also validate the robust and non-robust methods in the real data by adding extra noise to the diffusion signals. The mean of absolute value of differences in FA, MD and orientations are shown between tensors estimated from noisy data and raw data.

In Chapter 3, we focus on multiple fibres problems. Some recent techniques of multiple tensors models are reviewed. As the information of neighbouring tensors is greatly influential when estimating multiple tensors in the intersection and has been ignored in the previously mentioned methods, we develop Bayesian single tensor model and Bayesian double tensor model which take into account the information of neighbouring tensors. The background of Bayesian statistics is introduced. In Bayesian models, the priors use the information of neighbouring tensors, and the variance is represented by the inverse Gamma distribution. The likelihood function and posterior distribution of single and double tensor models have been deduced. Several indices of a model selection criterion are stated. We also proposed

Bayesian neighbour algorithm to estimate the single and double diffusion tensors with model selection available. The starting point and the phase portrait have been discussed and the procedure of the algorithm is stated in the framework. In the simulation studies, we simulate tensors fields that consist of single and double tensors at the crossing regions. In the first three simulation studies, we examine the dependence of the estimation on the FA and MD of two simulated tensors and angle between the two principal diffusion orientations. The last simulation study is focused on measuring the goodness of fit of the estimated double tensors. A real DTI image from a healthy human brain in coronal view is presented. The region of pons and corpus callosum is enlarged and the Bayesian models and BN algorithm are performed in the real data. Extra Gaussian noise is added to such DTI image. We estimate the double tensors again and compare the results between raw data and noisy data.

In Chapter 4, the estimation will be considered the positive semi-definite symmetric property of tensors in the tensor field. We define the median covariance tensors in terms of different Euclidean and non-Euclidean metrics. The non-Euclidean estimation techniques include power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes metrics. The tensor estimators are parameterised when necessary to ensure the symmetric semi-positive definiteness property. We also consider the weighting functions and define the weighted median covariance tensors in terms of various metrics. The geodesic anisotropy and procrustes anisotropy have been compared with fractional anisotropy index. The weighted non-Euclidean median tensor estimators are compared with Euclidean mean estimators in simulation studies and real data in terms of the determinant, MD, FA and orientation. The comparison between different median tensor estimators is made from a single covariance tensor estimation to the tensor field smoothing and interpolation. The non-Euclidean distances between each two estimated tensors are also compared. The non-Euclidean median tensor estimators are used to smooth and interpolate the tensor field with weighting functions in simulation studies and real data. We also validate our various non-Euclidean metrics. The root-mean-square deviation of FA, determinant and orientations of tensors are compared in the corpus callosum region. Extra Gaussian noise is added to the diffusion signals. The Euclidean mean and various non-Euclidean median tensors are used to smooth the tensor field. We compare the results with smoothing the original raw tensor field.

Finally, the conclusion of the research and the future work will be summarised in Chapter 5.

Chapter 2

Robust Algorithm For Tensor Estimation

2.1 Introduction

Diffusion Tensor Imaging (DTI) is an advanced method of Magnetics Resonance Imaging (MRI). It displays an insight structure of the biological tissue in the human body *in vivo* by directionally describing the water molecular diffusion [15, 7]. A diffusion tensor, a covariance matrix related to the molecular displacement at a particular voxel in the brain, is in the non-Euclidean space of 3×3 positive semi-definite symmetric matrices. The diffusion tensor is estimated at each voxel in the brain and is generated by fitting a physically motivated model on measurements from the Fourier transform of the molecule displacements [3]. However, the signal variability of diffusion tensor imaging is influenced by various contributions: thermal noise, spatially and acquisition-related factors [19], which will be explained in the next section. The outliers occurred in the measurement and estimation of DTI procedure, may affect the accuracy of tensor estimation. Traditional least squares methods take into account signal variability produced by thermal noise but ignore other causes, for example, a cardiac pulsation, bulk head motion, respiratory motion, or even hardware imperfection may influence the results [71]. In such cases, the linear regression model shows conflict with outliers and other aberrations. The level of confidence intervals and tests can be weakened in presence of the noise arising in the measurement of diffusion attenuations. Since the commonly used least squares methods are not reliable in detecting the outliers, robust methods are recommended due to their high efficiency when handling these aberrations. In this chapter, we propose robust methods with the forward search algorithm to improve the tensor estimation procedure to reduce the effect of possible outliers to a desirable level.

2.2 Noise and artifacts

We explain the noise or artifacts arising in the measurement of DTI signals, including subject motion, eddy currents, magnetic susceptibility effects, image noise and hardware imperfections.

Subject motion

The subject motion may result in ghosting or artifactually redistribution of signal intensities during the MR-DTI experiment. Due to rigid body motion, artifacts can be corrected by applying a uniform phase to an entire image [7]. In contrast, other physiological motions, for example, eye movements or pulsation of cerebrospinal fluid, may not correct by the current model.

Eddy currents

In DTI scanning, the gradient coils produce rapidly switched magnetic field gradients. Additional eddy currents are induced in the electrically conductive structure of DTI machine. This results in unwanted and slowly decaying magnetic fields. Thus, the gradient field is different from our preset and prescribe gradient field, resulting a different $b-matrix$. In addition, there is an geometrical distortion during the readout of the DTI image. As each voxel is calculated using prescribed gradient coefficients, there is a misregistration of data. Especially on single-shot echo-planar image, it is easily affect by eddy currents. To reduce this artifact, bipolar diffusion-encoding gradients [2] to diffusion-encoding gradients can be applied at the acquisition stage. A new tool 'FSL-eddy' is used to correct the eddy currents and movements in diffusion data [65].

Magnetic susceptibility effects

Local magnetic field gradients may be produced during large discontinuities in bulk magnetic susceptibility, especially in echo-planar imaging. These local gradients are easily confused as diffusion gradients, thereby $b-matix$ may spatially vary. This problem is more severe under a strong magnetic field (3 Tesla and above) [7].

Image noise

The background noise can bias the eigenvalues of a tensor, **D**. It results in anisotropic images from isotropic tensors and makes anisotropic tensors much more anisotropic [82].

The bias may increase the difficulty of the calculations by tending to a higher order measures. This problem can be improved by the lattice Anisotropy index proposed by Pierpaoli and Basser [82]. The RF noise also induces a significant bias in the mean and variance of the eigenvectors of **D** [9], but there are no remedies to correct it. Negative eigenvalues of diffusion tensor may present by the noise with current linear and non-linear regressions. As the diffusion tensor must be positive semi-definite, or non-negative, some constraint of positive definiteness can be added on the tensor.

Hardware issues

In the DTI scanning, additional signal attenuation is presented if the magnetic field is not properly shimmed. It is required to measure the background directly and incorporate them explicitly in the $b-matix$ [47]. Non-linear and calibrated gradients can lead to miscalculation of diffusion coefficients.

2.3 Robust statistics

Mathematicians have an awareness that some of the most general statistical regressions are excessively sensitive to small deviations from assumptions [42]. The term 'robust' was first coined by George Box in 1953, and robustness theory has developed in 1960s [38]. Lately, Peter J. Huber's achievements, from his fundamental paper about robust estimation [43] to his Wald lectures notes [40, 41], are the milestone of robust theory. The study of influence functions in robust estimation is referred to in Frank Hampel [36, 37] and the various robust estimators and generalisations are summaries in Huber's paper [39]. For more materials of mathematical views of robustness, it referred to [42, 44]. The robust statistics shows an alternative solution for the optimal procedures which the classical statistics fails to achieve in instability problems.

The robust procedure should satisfy the following features:

- Consistency, asymptotic normality and high relative efficiency of the estimators;
- Stability, relative insensitivity to slight violations of the model;
- Simplicity of the theory and ease of computation.

Some robust statistics also seek to achieve the highest possible breakdown point, which will be explained in the next section.

2.3.1 Robust indices

Consider a sample S consisting of n data, and let T be a regression estimator. If all possible corrupted samples S_c that contains any m of the replacing original data by arbitrary values, The maximum bias that can be caused by such an operation is

$$bias(m;T,S) = \sup_{S_c} \|T(S_c) - T(S)\| \tag{2.1}$$

where the supremum is over all possible S_c. It is a non-decreasing function of m. If the bias is infinite, m outliers can have an arbitrarily large effect on T. Therefore, the *breakdown point* of the estimator T at the sample S is defined as

$$\varepsilon^*(T,S) = \min\left\{\frac{m}{n}; bias(m;T,S)\right\} \tag{2.2}$$

The *breakdown point* of an estimator is the proportion of observations that an estimator can handle before leading to incorrect outlier detection [90, 42]. The highest possible breakdown point level is 50%. An estimator with a breakdown point approximately equal to 50% is called a *high breakdown* point estimator. The mean has 0% breakdown point, and the median has 50% breakdown point.

The robust measurement of the spread of the sample is given by the *median of all absolute distances*(MAD) from the sample median:

$$\text{MAD} = \underset{j=1,\cdots,n}{\text{median}} \left\| x_j - \underset{i=1,\cdots,n}{\text{median}}(x_i) \right\| \tag{2.3}$$

In order to use this robust estimator in the usual scale of Gaussian distributions, it needs to take a correction factor 1.4826 [91], which is approximately equal to $1/(\Phi^{-1}(3/4))$, Φ^{-1} is the inverse of the normal cumulative distribution function. s is approximate to the population standard deviation for normally distributed large samples:

$$s = 1.4826\text{MAD} \tag{2.4}$$

We then use the standardized values to detect the outliers:

$$z_i = \frac{x_i - M}{s} \tag{2.5}$$

M is the sample median. The z_i scores are compared to some cut-off values. Cut off values of 3 or more are usually considered very conservative, 2.5 or even less moderately

conservative [73, 66]. Since the difference of the stringency of the researcher's criteria, the cut-off values are defined and adjusted by researchers.

2.3.2 Robust estimators

The simplest robust estimation method is the least median squares method (LMS) [89]:

$$\text{minimize} \underset{i}{\text{median}} \sum_{i=1}^{N} r_i^2 \tag{2.6}$$

In the multiple regression model, we need to consider the residuals r_i:

$$r_i = y_i - \hat{y}_i \tag{2.7}$$

The LMS regression can achieve 50% breakdown point level. The LMS estimator continues to explain the majority of properties of the data even if almost half of the data are corrupted. In order to decide whether a residual from a robust regression is unusually large, we have to set a boundary. For this purpose, we need to estimate the spread of the error terms. Therefore, the scale estimate in LMS regression is

$$\hat{\sigma} = 1.4826 \sqrt{\underset{i=1,\cdots,n}{\text{median}} r_i^2} \tag{2.8}$$

Note LMS does not throw away 50% of the data. Rather, it finds a fit to the majority of the points, which can then be used to identify the actual outliers.

Rousseeuw [89] also proposed the least trimmed squares (LTS) estimator. Let $r_{(1)}, \cdots, r_{(N)}$ be the ordered squared residuals of the estimator from the smallest to largest, i.e. $r_{(1)}^2 \leq r_{(2)}^2 \leq \cdots \leq r_{(n)}^2$. The LTS estimator is given by:

$$\min \sum_{i=1}^{h} r_{(i)}^2 \tag{2.9}$$

where h is an positive integer and $h \approx \frac{n}{2}$.

The LTS estimator is also a high robustness estimator and its breakdown point is approximate to $\frac{1}{2}$ (if h is the right fraction of n). The high breakdown point of LTS estimator means that it can consist of more than 50% of that data remain correctly. On the other hand, if a model only contains normal errors, the LMS and LTS estimators is not highly efficient.

Comparing theses two methods, tt has been shown the LTS estimate converges much faster than the LMS estimate, this is not the case of the regression that is non-linear however [93].

There are numerous other robust estimators as well. The *M-estimator* was introduced by Huber for location model [43] and extended to the regression model later [41]. M stands for maximum likelihood-type [42]. For a finite sample $x_1, \cdots, x_n$, let ρ be an even function with a minimum at zero and close to the assumed distribution, the *M-estimator* θ is

$$\hat{\theta} = \arg\min \left(\sum_{i=1}^{n} \rho(x_i, \theta) \right) \tag{2.10}$$

Moreover, more generalised M-estimators (GM-estimators) were studied by Mallows and Schweppe. The breakdown point of GM-estimators decreases when the dimension of the estimators increases.

The *S-estimator* is based on the minimisation of the dispersion of the residuals $r_1, \cdots, r_n$ of an estimator θ:

$$\hat{S} = \arg\min s(r_1(\theta), \cdots, r_n(\theta)) \tag{2.11}$$

The dispersion of $s(r_1(\theta), \cdots, r_N(\theta))$ is defined as the solution of following equation for a given K:

$$\frac{1}{n} = \sum_{i=1}^{n} \rho(\frac{r_i}{s}) = K \tag{2.12}$$

where the function ρ is symmetric and continuously differentiable with $\rho(0) = 0$.

Further details of the dispersion function are referred to in [89].

The *repeated median* [97] has a 50% breakdown point, but it adds the complexity on computing tensor data.

$$\hat{\beta}_j = \underset{i_1}{\text{median}} \left(\underset{i_2}{\text{median}} \left(\cdots \left(\underset{i_n}{\text{median}}\, \beta_j(i_1, i_2, \cdots, i_n) \right) \cdots \right) \right) \tag{2.13}$$

However, most of these estimators suffer from a low breakdown point level, which means they have a low tolerance for outlier detection [90, 71], including the median of pairwise slopes, R-estimators [45] and L-estimators, etc.

2.4 Methodologies

The least squares method(LS) dates back to Gauss and Legendre, around 1800 [102]. It finds the estimator by minimising the residuals

$$\text{minimize} \sum_{i=1}^{n} \varepsilon_i^2, i = 1, \cdots, n \tag{2.14}$$

2.4.1 Least median squares (LMS) regression

The simplest very robust estimation method is the least median squares (LMS) approach. This regression is introduced by Rousseeuw and has the highest breakdown point level, 50% [89]. The LMS estimate minimises the median of squared residuals ε_i and is expressed as follows:

$$\text{minimize} \underset{i}{\text{median}} \ \varepsilon_i^2 \tag{2.15}$$

In the multiple regression model, we need to consider the residuals ε_i:

$$\varepsilon_i = y_i - \hat{y}_i \tag{2.16}$$

Thus in DTI, the objective function to be minimized is: LMS

$$f(\beta_{LMS}) = \underset{i}{\text{median}} \left[S_i - S_0 \exp \left(\sum_{j=1}^{N} X_{ij} \beta_j \right) \right] \tag{2.17}$$

The Cholesky reparameterisations can be also added in the minimisation with the diffusion elements:

$$f(\rho_{CLMS}) = \underset{i}{\text{median}} \left[S_i - S_0 \exp \left(\sum_{j=1}^{6} X_{ij} \beta(\rho)_j \right) \right] \tag{2.18}$$

2.4.2 Least trimmed squares (LTS) regression

Another very robust estimator with an asymptotic breakdown point of 50% is the least trimmed squares(LTS) estimator [89]. The LTS regression minimises the sum of the hth smallest squared residuals. It is given by

$$\text{minimize} \sum_{(i)=1}^{h} \varepsilon_{(i)}^2 \tag{2.19}$$

where $\varepsilon^2_{(1)} \leq \varepsilon^2_{(2)} \leq \cdots \leq \varepsilon^2_{(n)}$ are the ordered squared residuals and $h \approx \frac{n}{2}$.

Different value of h gives different trimming proportion [89]. Rousseeuw and Hubert suggest $h = (n+p)/2$ or $h = (n+p+1)/2$ (whichever is an integer) [92], where p is the number of unknown parameters. β_{LTS} has a convergence rate of $n^{-1/2}$, is better than β_{LMS} with a convergence rate of $n^{-1/3}$ for large samples [89, 103].

To model the diffusion model in DTI, the objective function is required to be minimised:

$$f(\beta_{LTS}) = \sum_{(i)=1}^{h} \left[S_i - S_0 \exp\left(\sum_{j=1}^{N} X_{ij}\beta_j \right) \right]^2 \tag{2.20}$$

2.4.3 Forward search (FS) algorithm for robust estimations

The forward search (FS) algorithm is introduced by Riani and Atkinson [86], an efficient algorithm for multivariate data with excellent power and type one error rate [88]. The forward search starts from a subset of size $m = p$ of the n observations, where p is the number of unknown parameters. The least median squares method is applied to find estimators of regression parameters as a starting point. The result is identical to the least trimmed squares approach when $n >> p$ [87]. The very robust estimates ensure that the observations are the central part of a sample. Let $S^{(m)}$ be the optimum subset of size m, which yields the estimators $\hat{D}^{(m)}$. The residuals ε_i are calculated from the n observations:

$$\varepsilon_i = S_i - S_0 \exp(-b g_i^T \hat{D}^{(m)} g_i), \qquad \text{with } i = 1, \cdots, n. \tag{2.21}$$

The likelihood ratio test is used to see whether an observation is an element in $S^{(m)}$. The deletion residual for $n-m$ observations is:

$$\frac{\varepsilon_i}{s^{(m)}\sqrt{1+\frac{1}{m}}}, \qquad i \notin S^{(m)} \tag{2.22}$$

where $s^{(m)}$ is the estimator of absolute mean error from truncated samples of observation m.

Denote the observation i_{min} with minimum absolute deletion residual. The criterion $\varepsilon_{i_{min}}$ is used to determine whether the observation i_{min} is an outlier. If it is an outlier, the process stops here and the $n-m$ observations are treated as extreme points. $\hat{D}^{(m)}$ is the optimum estimator required. Otherwise, the subset $S^{(m+1)}$ is formed by $S^{(m)}$ and $S_{i_{min}}$. It consists of the $m+1$ observations with the smallest absolute values of residuals ε_i. Applying the linear

least squares method on these $m+1$ observations, a new estimator is obtained:

$$\hat{D}^{(m+1)} = X^{(m+1)} y^{(m+1)} \tag{2.23}$$

where $X^{(m+1)}$ is an $(m+1) \times p$ full-rank matrix of known constants and

$$y^{(m+1)} = [\log(S_i/S_0), \cdots, \log(S_{m+1}/S_0)]^T \tag{2.24}$$

The search moves forward with the new subset $S^{(m+2)}$ after calculating the residuals by substituting $\hat{D}^{(m+1)}$ in Equation (2.21). The size of this subset is increased by 1 each time until the observations not in the subset are recognised as outliers. The FS algorithm is as follows:

Algorithm for FS

1. Initial=D_{LMS}, size $m = p$;
2. Calculate residuals ε_i;
3. Order residuals $\varepsilon_{(i)}, \varepsilon_{(1)} < \varepsilon_{(2)} < \cdots < \varepsilon_{(n)}$, create $S^{(m+1)}$;
4. Estimate **D** in $S^{(m+1)}$ by CLMS or LTS;
5. Calculate the test statistics, if the criterion satisfied, stop the iteration, otherwise, go back to step 2.

2.4.4 Test criterion

Here are three methods to approximate the distributions of criterion $r_{i_{min}}$ on testing outliers [87].

Method 1. **Truncated samples**

A truncated normal distribution $r^z_{i_{min}}$ is simulated as in Equation (2.22). First, $m+1$ random numbers z_i are generated by

$$z_i = \Phi^{-1}(U_i) \tag{2.25}$$

where U_i belongs to an uniform distribution $[0.5-(m+1)/2n, 0.5+(m+1)/2n]$. Denote the most outlying observation $z_{i_{min}} = \min|z_i|$. Let $S_z^{(m)}$ be the sets of observations excluding $z_{i_{min}}$:

$$S_z^{(m)} = \{z_i | i \neq z_{i_{min}} = 1, \cdots, m+1\}. \tag{2.26}$$

Estimate the parameters the mean $\bar{z}^{(m)}$ and the standard deviation $s_z^{(m)}$ from $S_z^{(m)}$, thus the simulated value $r_{i_{min}}^z$ is acquired from

$$r_{i_{min}}^z = \frac{z_{i_{min}} - \bar{z}^{(m)}}{s_z^{(m)}\sqrt{(1+\frac{1}{m})}} \tag{2.27}$$

Repeat the processes until the distribution of the simulated value is derived.

The simulation study shows that distribution gains better regression by multiplying a correction factor for the upper percentage points:

$$r_{i_{min}}^{adj} = \sqrt{\frac{m+\theta p}{m}} \frac{z_{i_{min}} - \bar{z}^{(m)}}{s_z^{(m)}} \tag{2.28}$$

with $\theta = 0.7$. Empirically the adjusted statistics is futile for the lower percentage points.

Method 2. **Order observations**

There are n observations generating from a normal distribution just once before the forward search. Then the absolute values of residuals $e_{z_i}^{(n)} = |z_i - \bar{z}^{(n)}|$ from the mean is ordered. Each time, the subset $S^{(m)}$ contains the m smallest observations.

Method 3. **Order statistics**

The confidence level of Equation (2.27) can be derived from expected values of normal order statistics. A $100(1-\alpha)\%$ confidence interval for the outliers test is given by

$$\frac{\varsigma_{m+1,n} \pm \sigma_\varsigma \Phi^{-1}(\alpha/2)}{\sigma_T\sqrt{1+\frac{1}{m}}} \tag{2.29}$$

where

$$\varsigma_{m+1,n} = \Phi^{-1}\left\{\frac{8n+8m+7}{4(4n+1)}\right\} \tag{2.30}$$

$$\sigma_{\varsigma}^2 = \frac{(8m+5)(8n-8m-3)}{4n(4n+1)^2\ \phi^2\left\{\Phi^{-1}(\frac{8m+5}{2(4n+1)})\right\}} \tag{2.31}$$

and

$$\sigma_T^2 = 1 - \frac{2n}{m}\ \Phi^{-1}(\frac{n+m}{2n})\ \phi\left\{\Phi^{-1}(\frac{n+m}{2n})\right\} \tag{2.32}$$

2.5 Simulations

A simulation study is carried out to compare the properties of the estimates. We consider a problem of estimating a tensor **D** from linear least squares (LLS), cholesky non-linear least squares (CLNS), cholesky least median squares (CLMS) and least trimmed squares (LTS) regression from simulated diffusion signals. The signals S_i are generated as:

$$S_i = 100\exp(-bg^T D_n g) + \varepsilon_i, \varepsilon_i \sim N(0,5^2), i = 1,\cdots,32. \tag{2.33}$$

We examine three properties of the tensor: anisotropy, volume and orientation. The Fractional Anisotropy (FA) describes the fraction of the 'magnitude' of **D** that can be ascribed to anisotropic diffusion [15]. The Mean Diffusivity (MD), which measures the overall displacement of molecules, is the mean of eigenvalues. Recall the equation for FA (Equation (1.19)) is as follows:

$$FA = \frac{\sqrt{3[\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}}{\sqrt{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{2.34}$$

The determinant of **D** shows the volume of the tensor. The orientation is the principal direction of the diffusivity, which is characterised by v_1 of Equation 1.6.

Three different situations are taken account in this simulation study. In the first case, we consider three tensors D_1, D_2 and D_3 with the same size but different FA. The tensor D_1 is isotropic, i.e. the diffusivity is equal in every direction. The principal directions of D_2 and D_3 are the same. In the second case, two tensors D_4 and D_5 are considered. They have the same FA and principal direction, but the size is different ($|D_5| = 27|D_4|$). In the last case, we consider a series of tensors of different principal directions. D_6 is rotating through the following angles: 0°, 15°, 30°, 45°, 60°, 75° and 90°. The tensors D_n and its properties are shown in Figure 2.1 and Table 2.1.

In this research, we also add between 1 and 15 outliers to the signals S_i. Thus it shows the results in presence of outliers counting from 0% to 46.88% of diffusion signals S_i. The graphs illustrate the root-mean-square error of fractional anisotropic, determinant and angle between the D_n of estimators. The horizontal axis represents the number of outliers in S_i, the outliers are set as ten times of randomly selected signals. the results are based on 100 Monte Carlo simulations.

Figure 2.1 The tensors D_1,D_2, D_3, D_4 and D_6.

Tensor	Fractional Anisotropic	Determinant	Principal Direction
D_1	0	64	-
D_2	0.6163	64	[1 0 0]
D_3	0.8616	64	[1 0 0]
D_4	0.4629	6	[0.7660 0.6428 0]
D_5	0.4629	162	[0.7660 0.6428 0]
D_6	0.6030	3	[1 0 0]

Table 2.1 Fractional anisotropic, determinant and principal directions of D_1,D_2, D_3, D_4 and D_6.

(i) Different FA, same size and same orientation

In Figures 2.2, 2.3 and 2.4, D_1, D_2 and D_3 are used in the simulations. The RMSE of FA and MD of tensors D_1 and D_2 becomes greater when the number of corrupted data increases by LLS, CLMS and LTS mothods. The LTS estimate and the LLS estimate are better than the other two estimates for spherical or pie-shape tensors, while the situation reverses when estimating a linear shape tensor. The CNLS is problematic with regard to estimating the determinant of tensors. The robust estimates are better than traditional estimates. The CLMS estimator has slightly less RMSE of determinant than the LTS estimator. There is great variation in RMSE of orientation between estimators and defaulted tensor. The CLMS estimator is better at estimating a spherical tensor while the LTS and LLS estimators are better at illustrating a fusiform shape tensor.

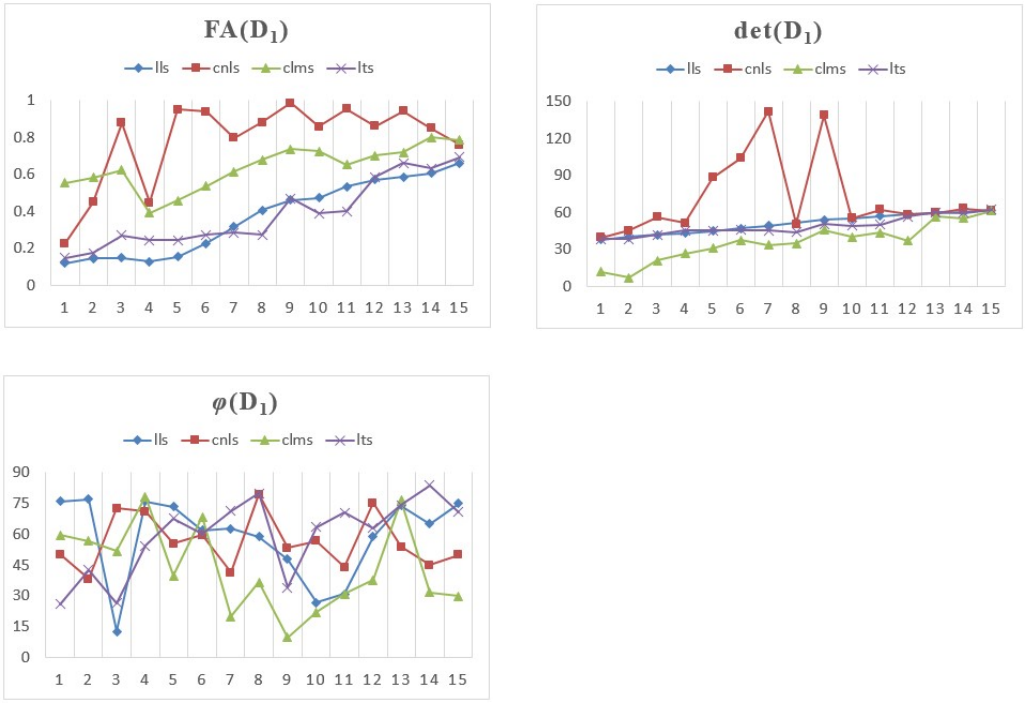


Figure 2.2 The Root-mean-square error of FA, determinant and angle of D_1.

(ii) Different size (determinant), same FA and same orientation

We compare two tensors: D_4 and D_5 in the Figures 2.5 and 2.6. As shown in the graphs, the techniques are better at estimating small tensors rather than large ones. The RMSE of LLS estimate is gradually increasing as the number of outliers increases. In contrast, the RMSE of the LTS estimate rises steeply when the number of outliers is greater than 11, one third of the signals. When the number of outliers becomes greater, the LMS and CNLS estimate begin to dominate the results. When the default determinant (D_5) is large, the RMSE of determinant and Angle is usually quite large. CLMS estimate is slightly better than the others in estimating the size of tensor. LTS and LLS have better estimators of FA.

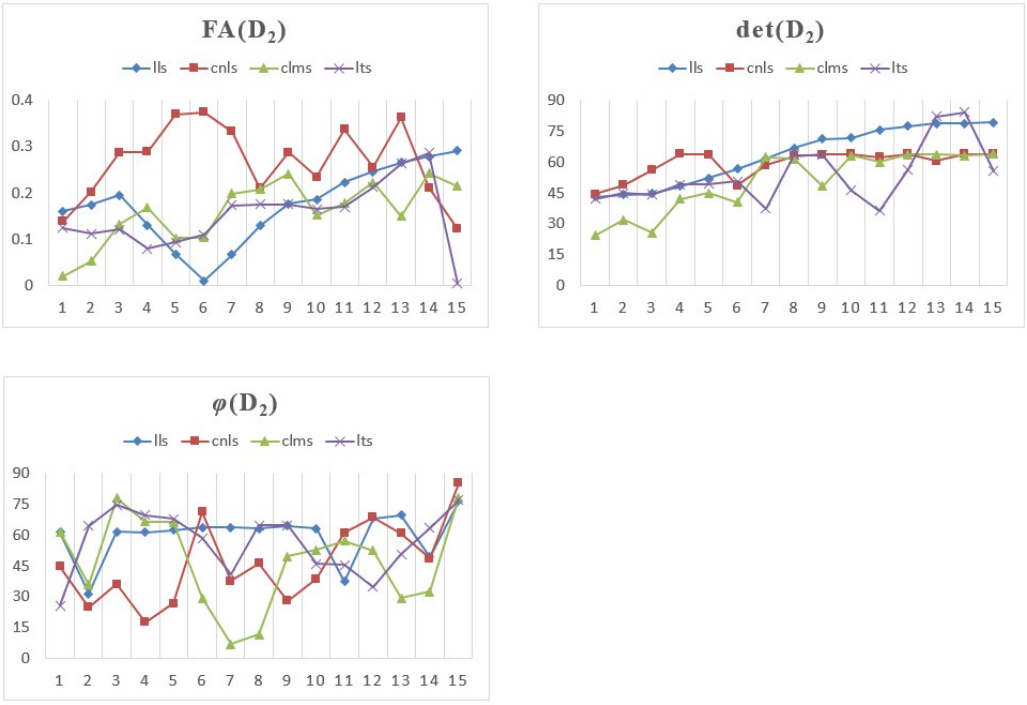


Figure 2.3 The Root-mean-square error of FA, determinant and angle of D_2.

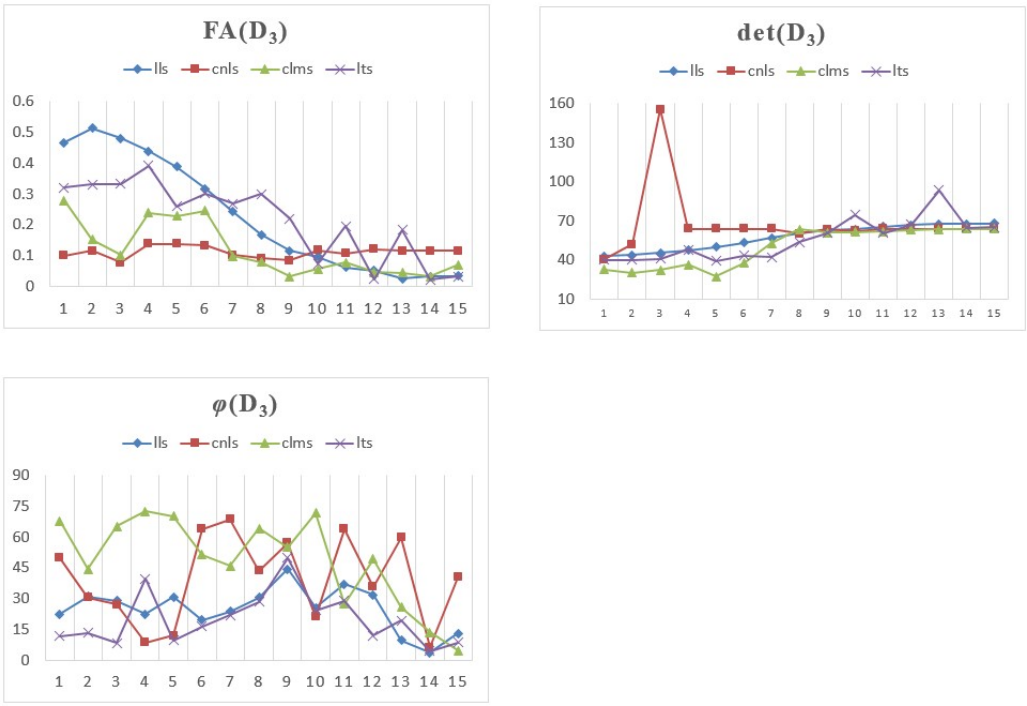


Figure 2.4 The Root-mean-square error of FA, determinant and angle of D_3.

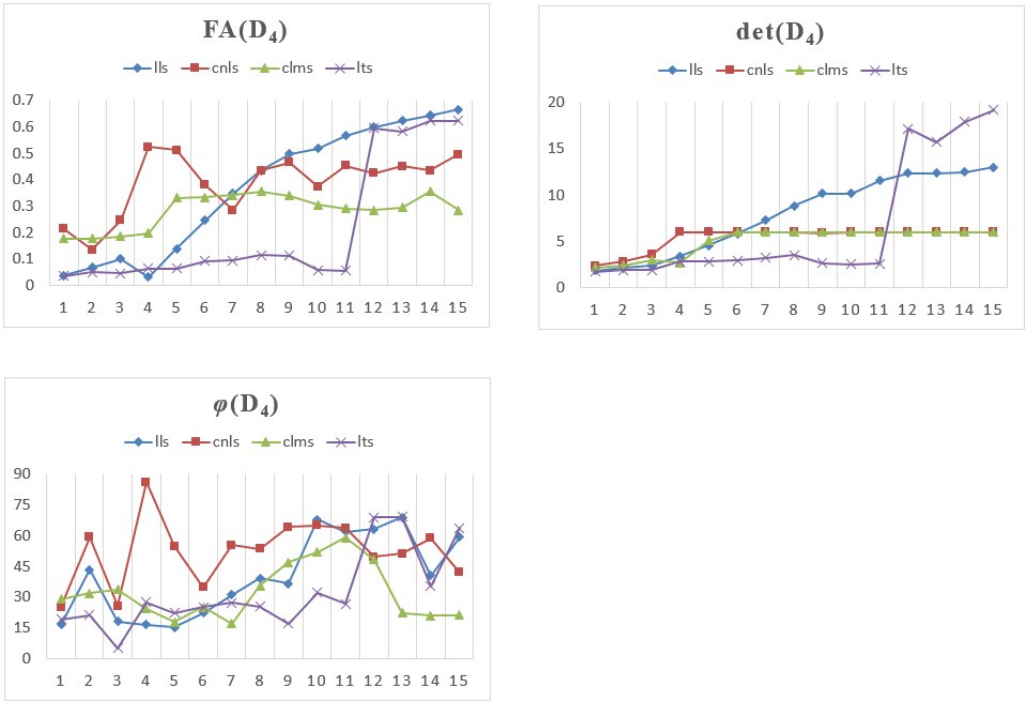


Figure 2.5 The Root-mean-square error of FA, determinant and angle of D_4.

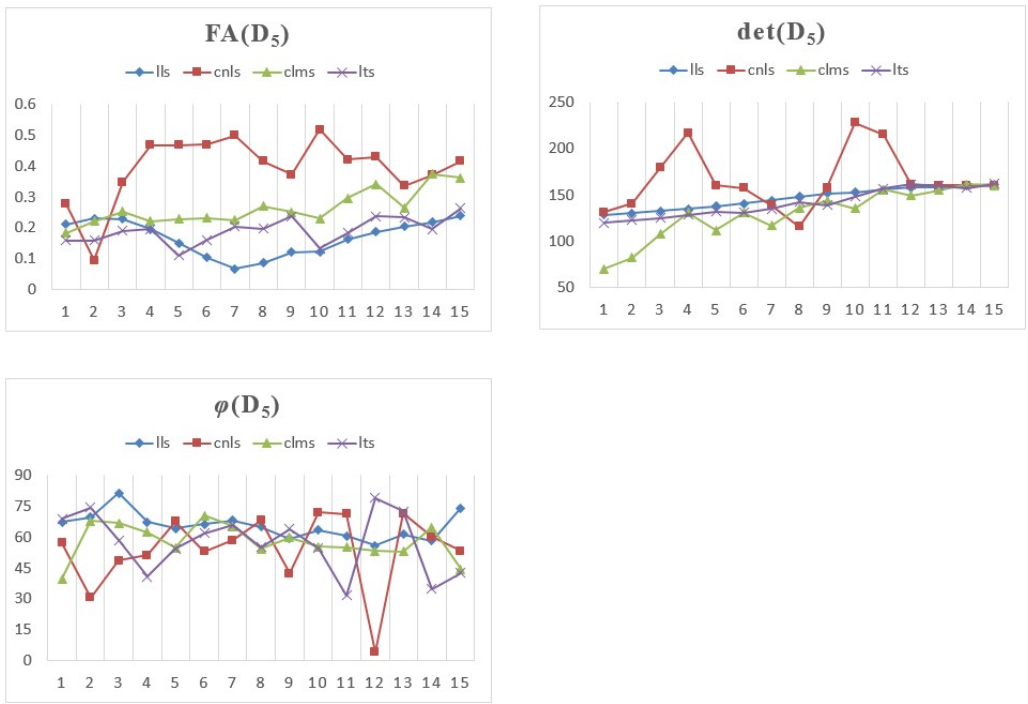


Figure 2.6 The Root-mean-square error of FA, determinant and angle of D_5.

(iii) Different orientation, same FA and same size

In this simulation, we examine the different orientation of an ovoid shape tensor D_6. The graphs of RMSE(FA) of the rotating tensor are similar in Figure 2.7. The CNLS estimator becomes unstable when the orientation gets larger and it has the worst results. The LTS estimate performs the best when outliers are less than one third of the signals, whilst the CLMS regression has the best results when there is more than one third of the data are outliers. In terms of the determinant, CNLS approach is still problematic. The LTS estimate obtains the best results when the number of outliers is small, while the CLMS estimate is best when there are more outliers. As the orientation of tensor D_6 gets bigger, the CNLS estimate has larger errors than other estimates. There is great variation in RMSE of angle in Figure 2.7. The RMSE of LTS estimate is slightly less than other three estimates and more stable.

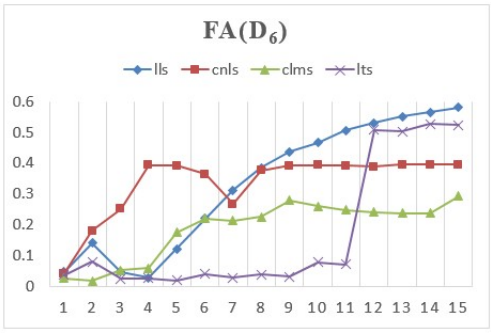
2.6 Real data analysis

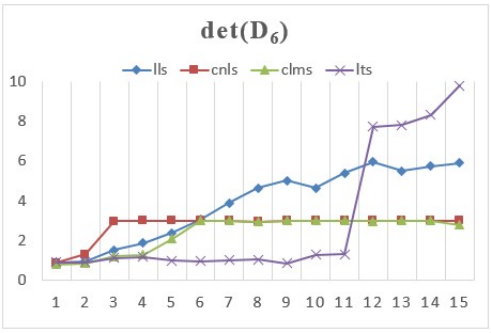
This study is based on the diffusion MR images from a healthy human brain. The current data contains $128 \times 128 \times 66$ voxels and each voxel has 32 diffusion signals $(S_1, S_2, \cdots, S_{32})$ under magnetics and one without (S_0).

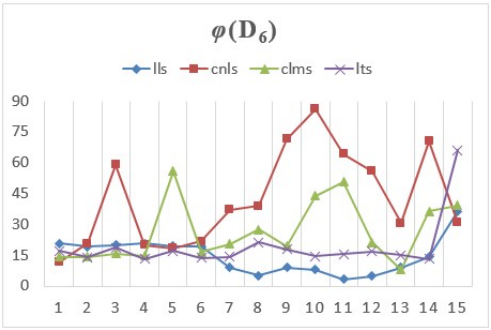
2.6.1 Robust estimation on a single tensor

The LLS, CNLS, CLMS and LTS methods are used to compute the tensors. A partial slice of the diffusion tensor FA image and its tensor field obtained by LLS method are shown in Figure 2.8. The corresponding tensor field is shown in Figure 2.9. The CLNS, CLMS and LTS methods produce very similar images as the LLS approach. Therefore, we enlarge four regions of tensors in Figure 2.9. Figures 2.10, 2.11, 2.12 and 2.13 describe the tensors estimated by four methods in the region (a), (b), (c) and (d) respectively. At each region, the central tensor or tensors will be estimated by these four different techniques, and the surrounding tensors remain the same.

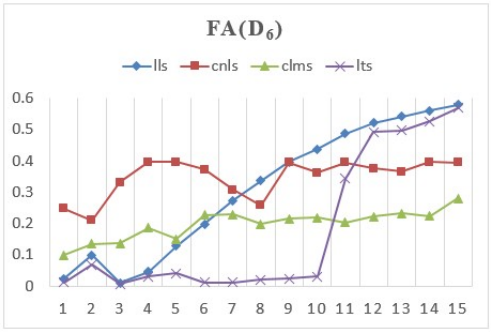
In the region (a), there are eight small volume tensors and a large size tensor at the bottom left corner. The orientation of LLS estimate is distinct with its neighbours, whilst the other three estimates coincide with the surrounding tensors in the Figure 2.10. The CLMS

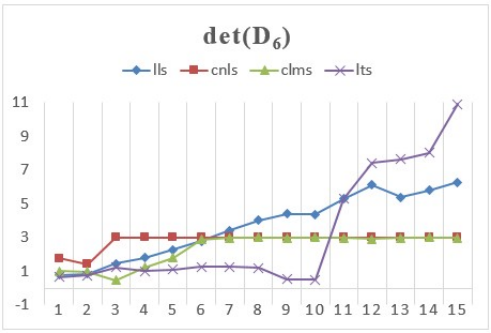


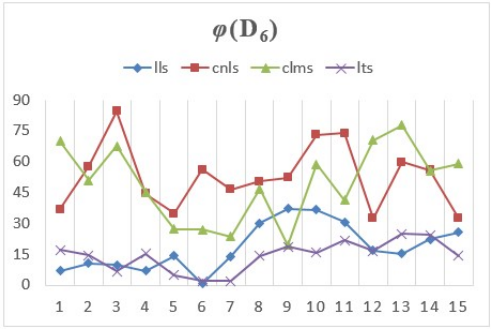




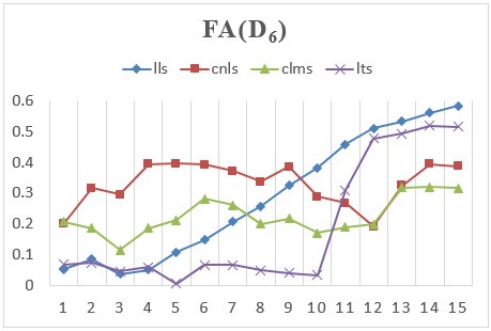
(a) 0 degree

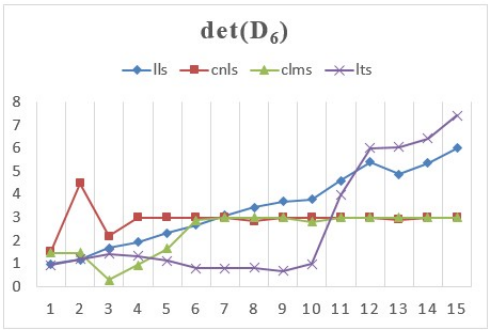


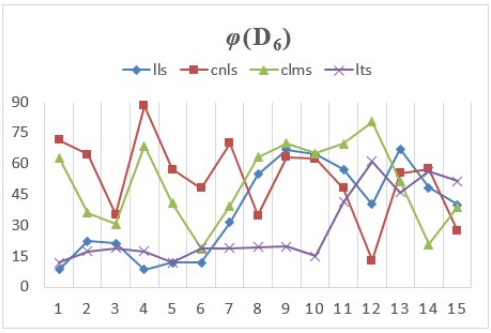




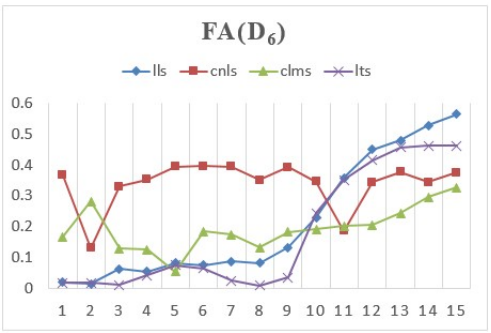
(b) 15 degree

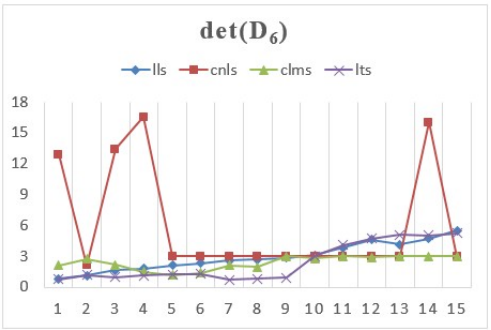


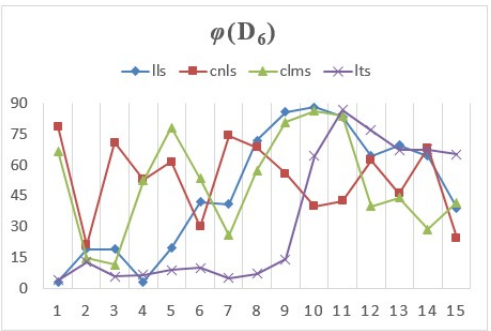




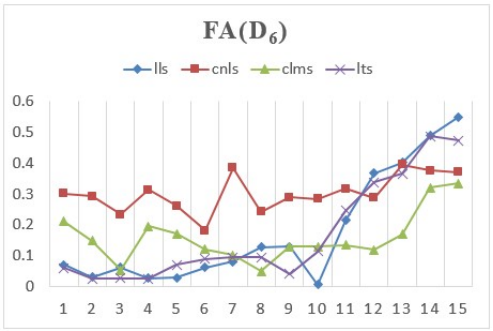
(c) 30 degree

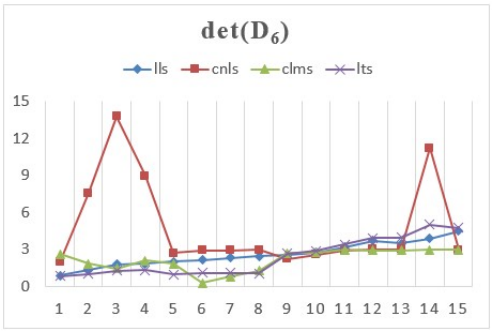


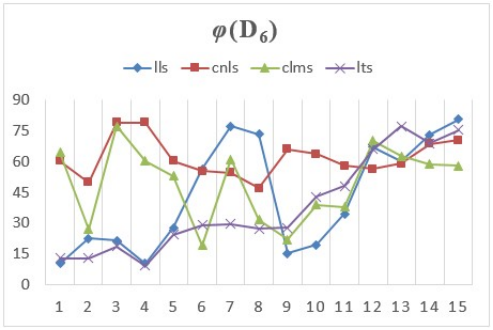




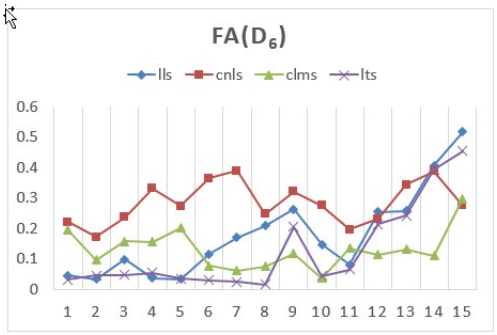
(d) 45 degree

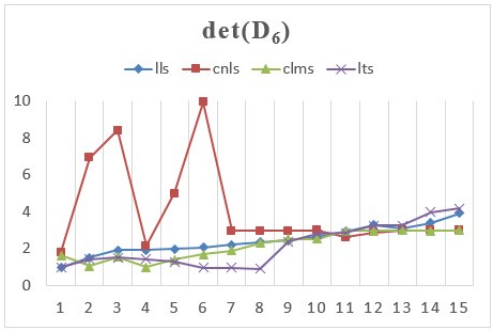


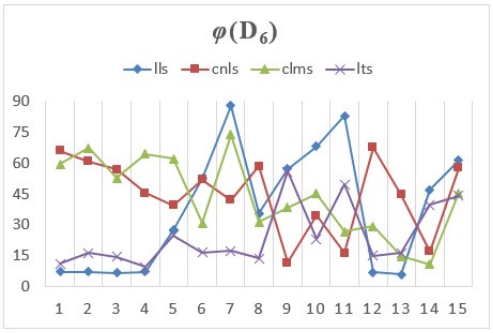




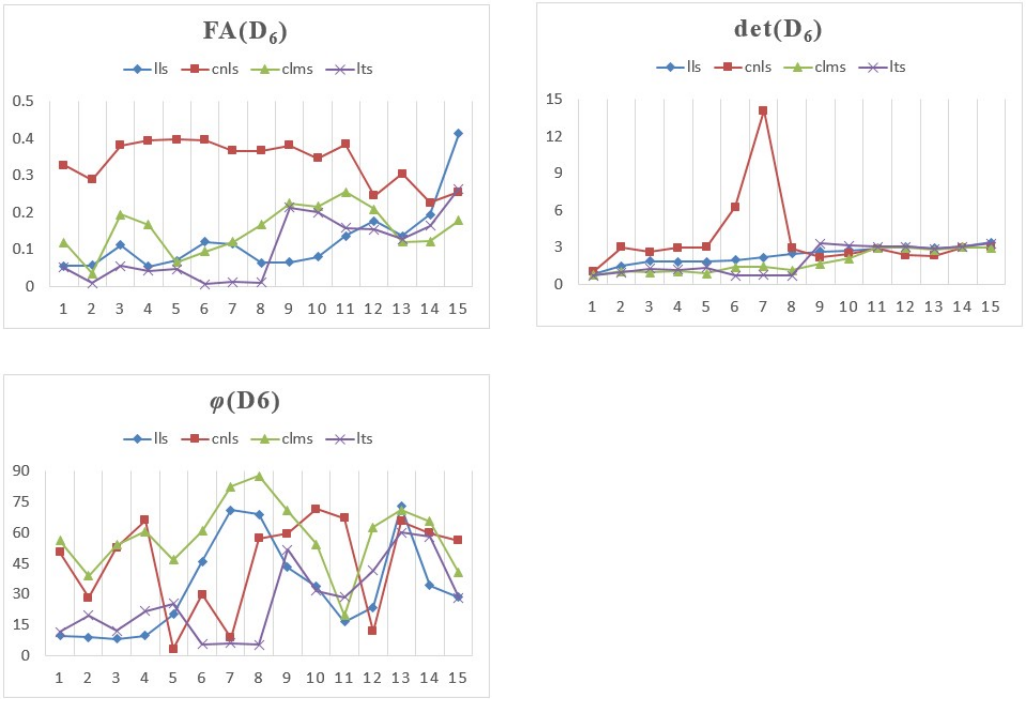
(c) 60 degree







(d) 75 degree



(g) 90 degree

Figure 2.7 The Root-mean-square error of FA, determinant and angle of D_6 and its rotations with angles 15°, 30°, 45°, 60°, 75° and 90°.

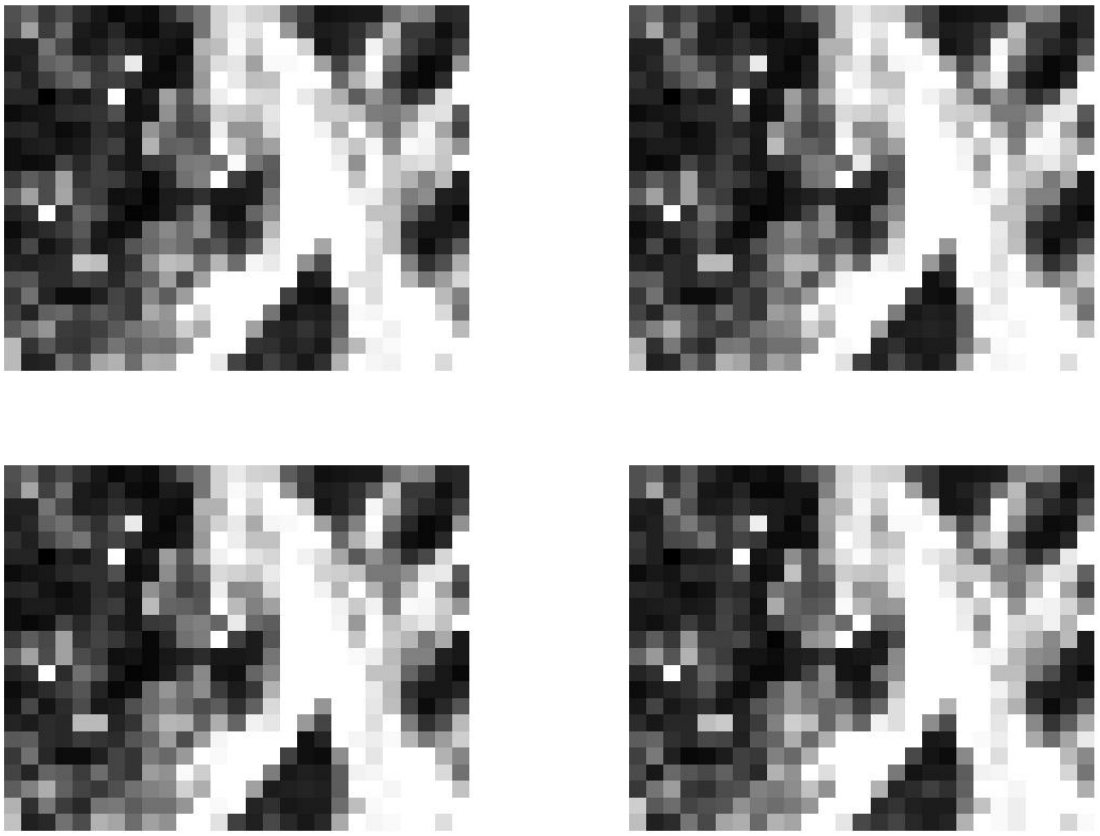


Figure 2.8 A real DTI FA map estimated by LLS (top left), CNLS (top right), CLMS (bottom left) and LTS (bottom right) methods.

approach has a tiny size, which is much smaller than its neighbour tensors. The CNLS and LTS methods obtain reasonable good results.

In Figure 2.11, the sizes of the tensors at the right and left column are quite large while the tensors at the central column are relatively small. The orientation of estimated tensor by LLS method is again totally different from its surrounding tensors. The CNLS approach results in a larger volume tensor. Although there are more large size tensors in the graphs, the upper and bottom tensors are small volume ellipsoid in the central column. The size of CLMS and LTS estimates are more appreciable than CNLS result.

The shape of the tensors in region (c) is more close to a sphere rather than an ellipsoid. The volume of CLMS estimate is very small in the Figure 2.12. Although it has a very small tensor at the top and a relatively large tensor at the bottom, the transforming between the

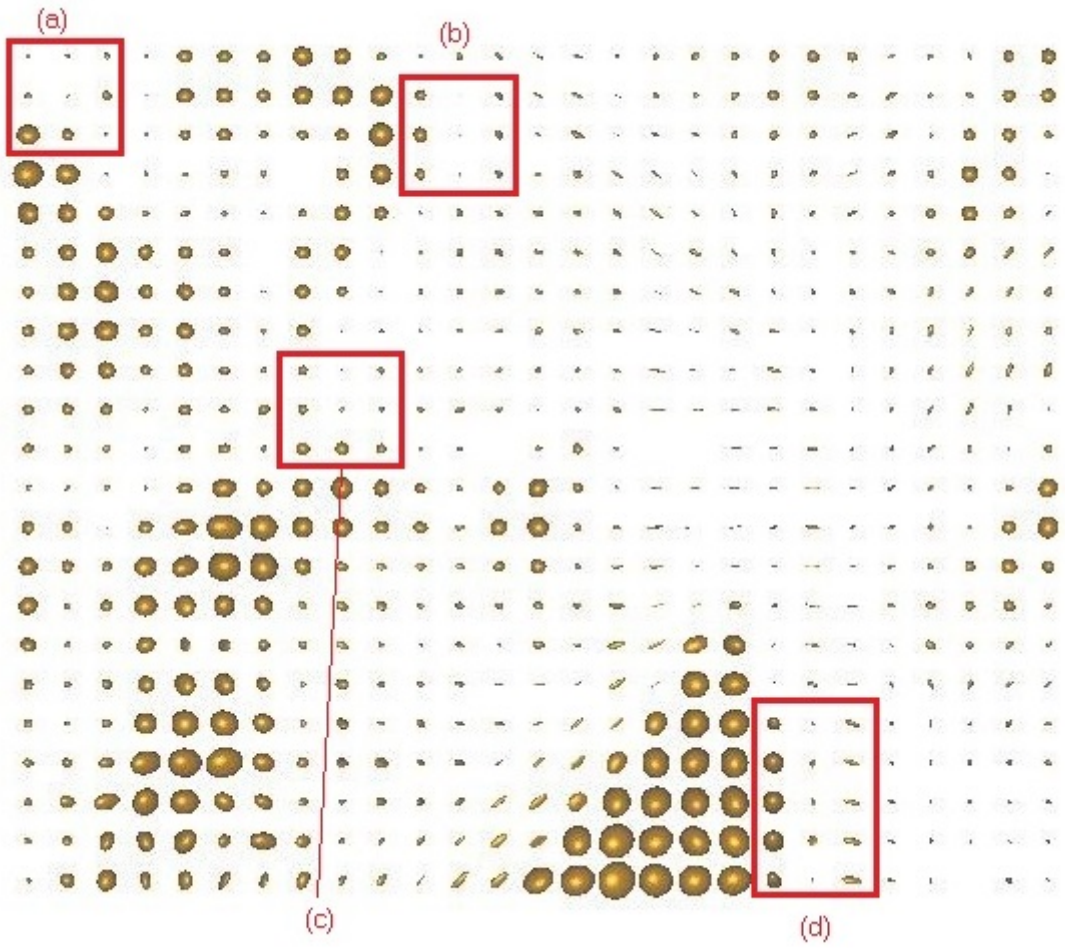
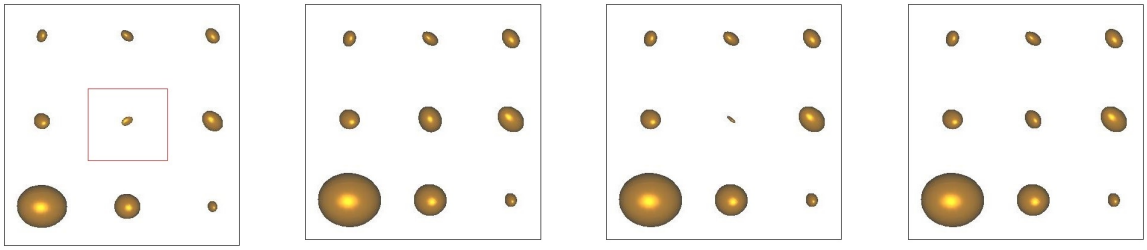
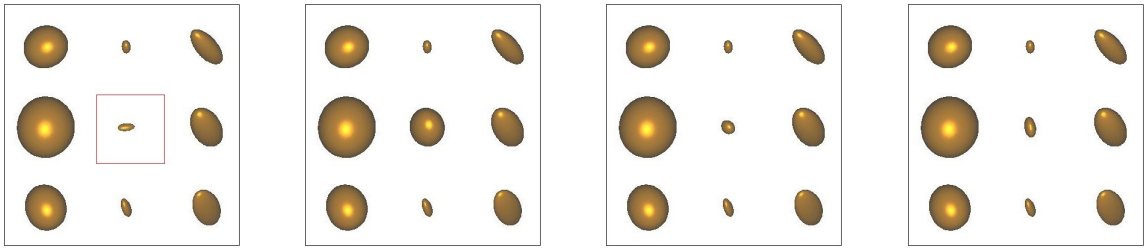


Figure 2.9 The tensor field of DTI image in Figure 2.8.



LLS CNLS CLMS LTS

Figure 2.10 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (a) of Figure 2.8.



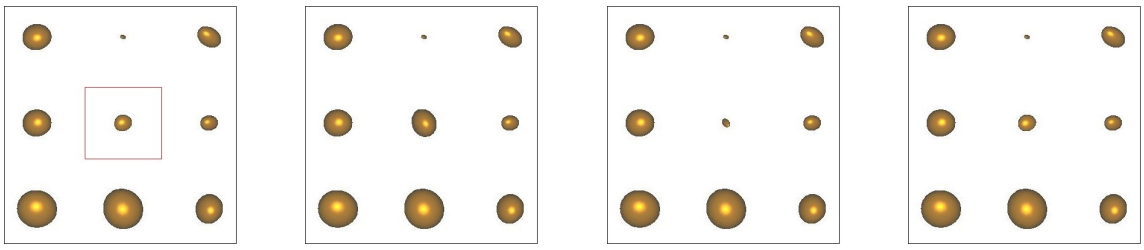
LLS CNLS CLMS LTS

Figure 2.11 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (b) of Figure 2.8.

central tensor and its bottom tensor is smoother than for the other approaches. The result of CLNS regression is more ellipsoidal. The LLS and LTS results are similar.

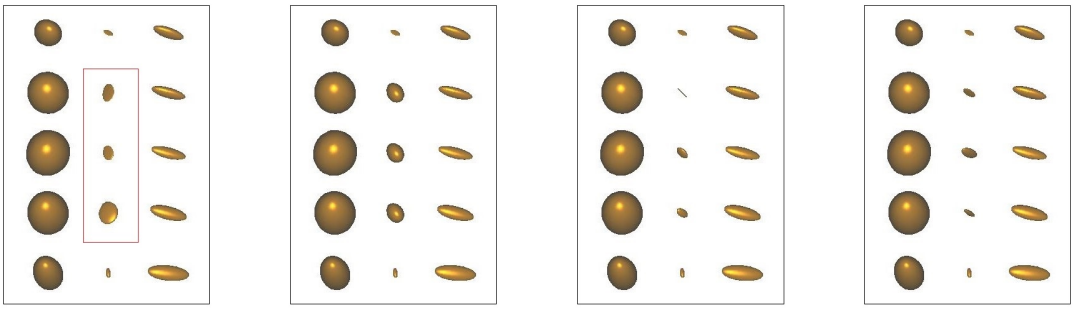
In the Figure 2.13, the central three tensor are generated by various techniques and the surrounding tensors are the same in the four graphs. The shapes of three tensors obtained by LLS method are disordered. The other estimates have similar orientations. The size of LLS and CNLS estimated tensors are the larger than CLMS and LTS approaches. Looking at region (d) in the Figure 2.8, there are all small size tensors in the column of the estimating tensors. Thus the robust methods are more preferable. Additionally, one of the three CLMS estimators has slightly different shape than the others. The results of LTS approach are more consistent in the figure.

We pick some tensors and compare them with their robust estimator. In the next section, we will exam the diffusion tensor by slices.



LLS CNLS CLMS LTS

Figure 2.12 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (c) of Figure 2.8.



LLS CNLS CLMS LTS

Figure 2.13 The tensors estimated by LLS, CNLS, CLMS and LTS methods with their surrounding tensors in region (d) of Figure 2.8.

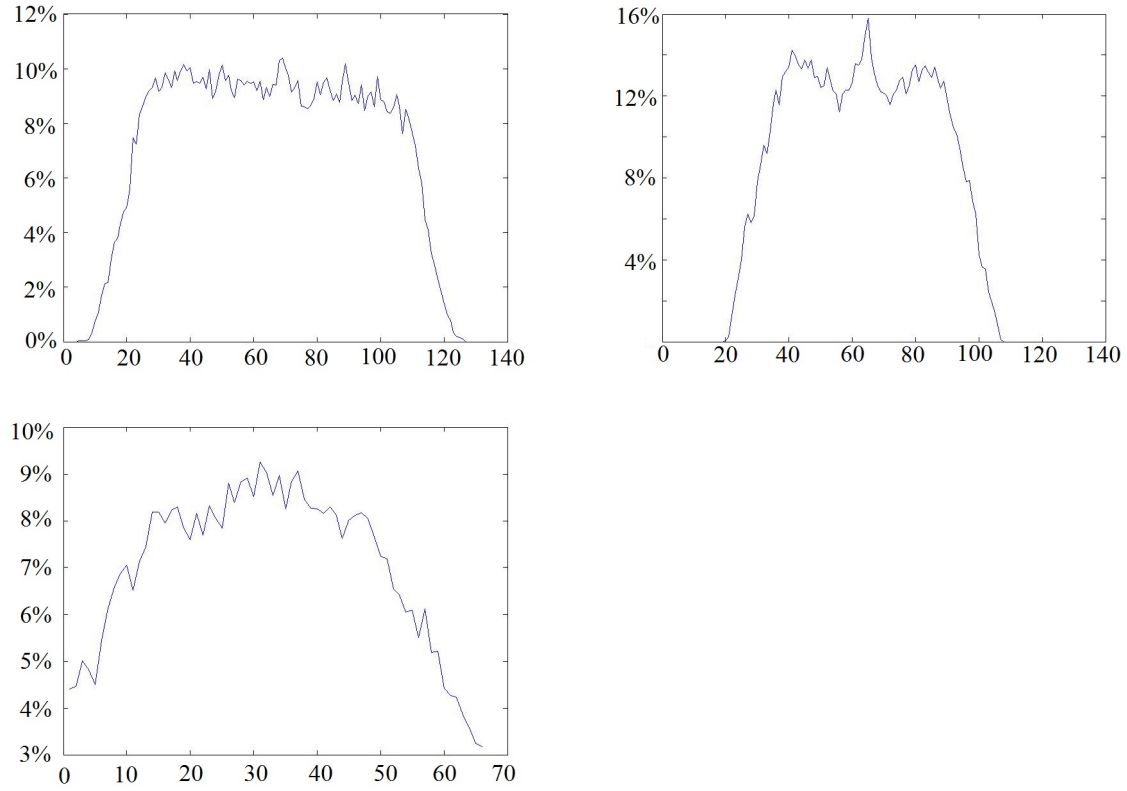


Figure 2.14 Percentages of voxels with outliers in each slice of DTI images from coronal, sagittal and transverse views.

2.6.2 Selection of outliers

First, we determine the outliers in the MR attenuations under magnetics with a conservative cut value of 3.0, i.e. if z_i (Equation (2.25)) is greater than 3, we consider the signal as an outlier. There are $75,964$ voxels considered as containing outliers, i.e. 7% of the whole data. Looking at each slice of MRI image from axial, sagittal and coronal views, we can see most outliers are in the central slices of the DTI images. There are $8,488$ voxels in each slice of coronal and sagittal view and $16,384$ voxels in each slice of transverse view. The number of voxels containing outliers in each slice of the DTI images from coronal, sagittal and transverse views is calculated. The percentages of these voxels are given in Figure 2.14, the horizontal axis represents the order of the slices and the vertical axis is the percentage of outliers voxels in terms of the whole voxels at each slice. Boxplots of these outliers voxels in each slice are also illustrated in Figure 2.15. The number of outliers in each voxel are summarised in Figure 2.16.

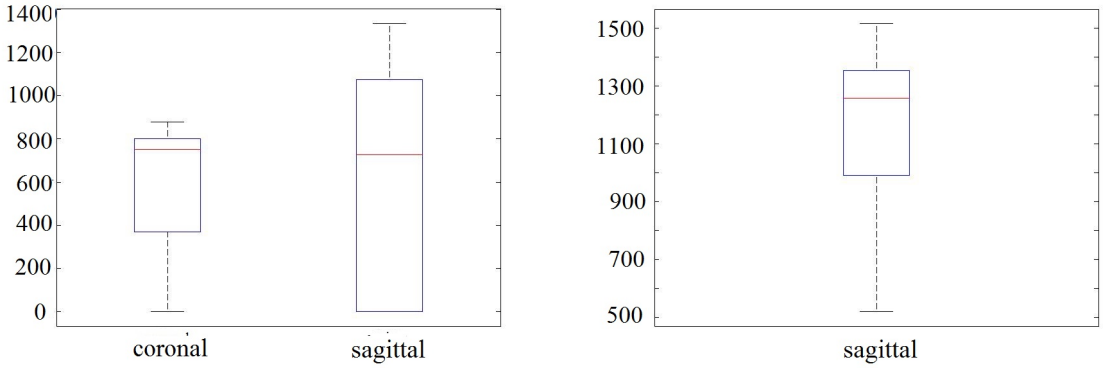


Figure 2.15 Box plot of number of voxels with outliers in each slice of DTI images from coronal, sagittal and transverse views.

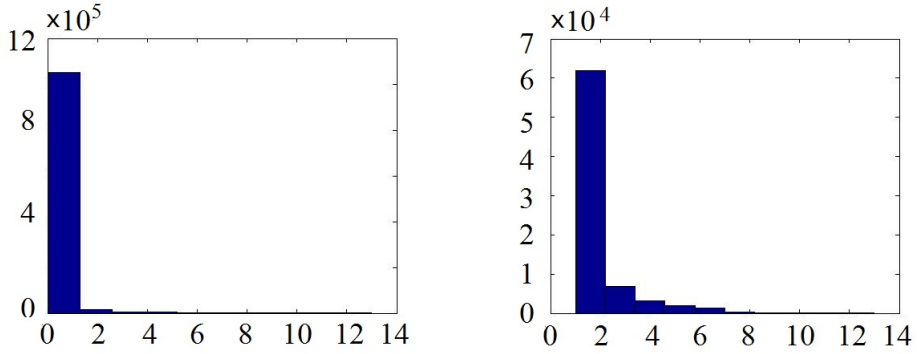


Figure 2.16 The number of outliers in each voxel of all voxels (left) and corrupted voxels (right).

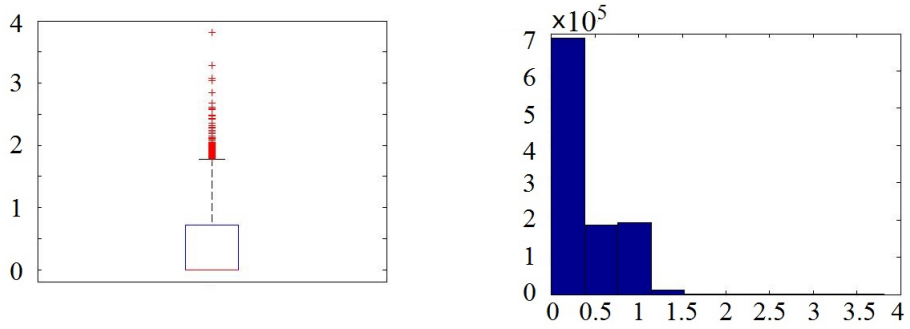


Figure 2.17 The box plot and histogram of $\frac{z_i}{32}$.

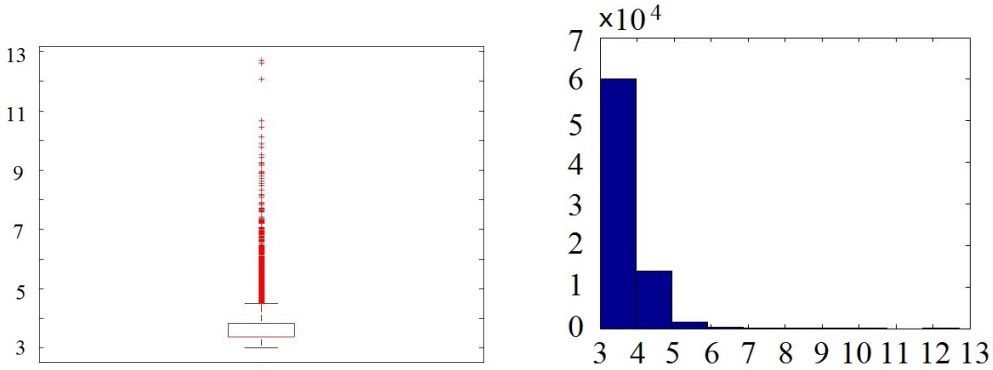


Figure 2.18 The box plot and histogram of $\bar{z}_i$.

We consider not only the number of outliers in each voxel but also the degree of these outliers voxels, in other words, the extent of the extreme value away from central majority values. The mean of z_i of all signals, i.e. $\frac{z_i}{32}$, ranges from 0 to 3.8081 with mean 0.7954. The mean value of z_i indicates that the cut-off value 3 is reasonable. The box plot and histogram of $\frac{z_i}{32}$ are shown in Figure 2.17.

Furthermore, we also compute the mean of the z_i of corrupted signals in each voxel, n is the number of corrupted signals in a voxel, $\frac{z_i}{32}$ is defined as follows:

$$\bar{z}_i = \frac{\sum(z_i)}{n}, \quad \text{for } z_i > 3. \tag{2.35}$$

The box plot and histogram of $\bar{z}_i$ are shown in Figure 2.18. The lower quartile of $\bar{z}_i$ is 3.3724. Let Ω denote voxels with $\bar{z}_i > 3.3724$ between the mean of voxels in Ω and those of which $z_i > 3$. A T-test is performed to examine the difference of Ω and the set of voxels containing z_i greater than 3. The null hypothesis is rejected at a significant level of 0.01, thus

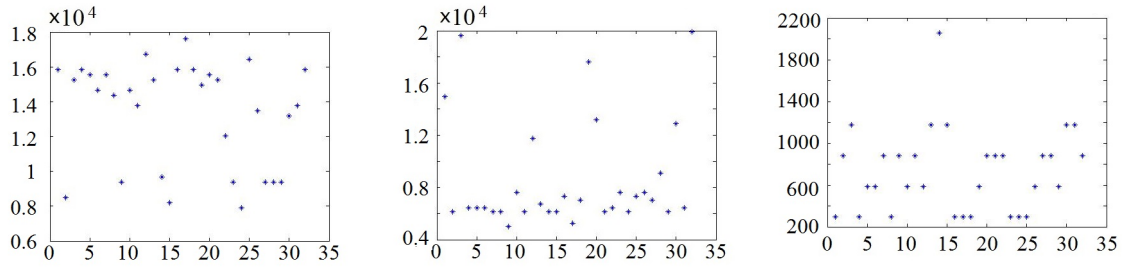


Figure 2.19 The signals under magnetics of voxel A, B and C.

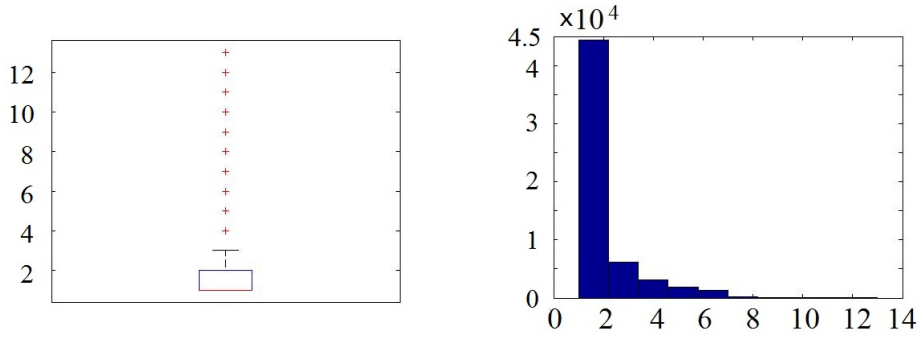


Figure 2.20 The box plot and histogram of number of outliers in each voxel in set Ω.

there is a strong difference between these two sets. We refer Ω to the voxels with outliers in our brain DTI image. There are $57,251$ voxels in the set Ω, nearly 5% of the whole data.

Why do we need to narrow the choice of outliers, not choosing the voxels containing z_i greater than 3? To explain this question, first, let's look at two voxels. Voxel A (at voxel coordinates $[23,49,23]$) has the majority signals around 1.6×10^4, while eight signals are around 0.8×10^4. Those eight signals have z_i values which are greater than 3. But the $\bar{z}_i$ of this voxel is 3.2881, not in Ω. The signals of such voxels which contain two level of values are likely to be the edge between two different tissues or the overlap of two or more fibres in the brain rather than outliers. The set Ω can exclude most of such voxels with two clusters of values because of their position in the brain rather than extreme values. As the mean number of outliers in Ω is 1.92, most voxels in this set have one or several extreme signals, such as voxel B (8 outliers) and C (1 outlier) in Figure 2.19. The histogram and box plot of the number of outliers in each voxel at Ω are illustrated in Figure 2.20. The properties of $\bar{z}_i$ in set Ω are described in Figure 2.21.

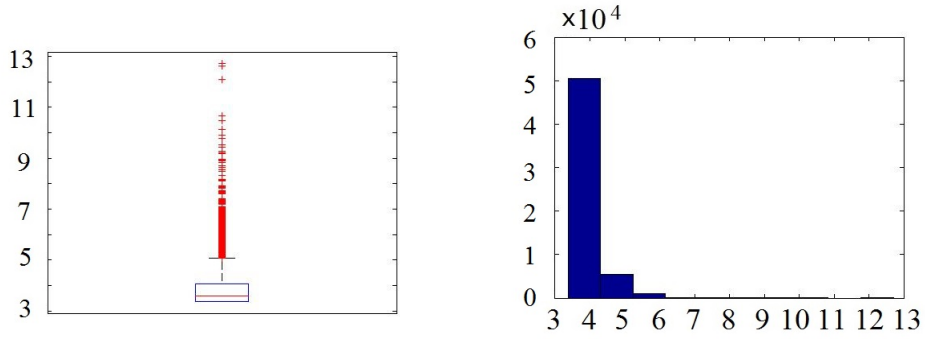


Figure 2.21 The box plot and histogram of $\bar{z}_i$ in set Ω.

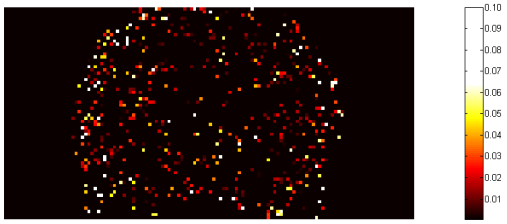
2.6.3 DTI slices from three dimensions

The real DTI images are estimated by four methods from three different views: coronal, sagittal and transverse views. A set of diffusion tensor imaging from a healthy human brain was provided by the Academic Radiology Department of Queen's Medical Centre, University of Nottingham. All the figures and graphs are obtained by the programme MATLAB(R2013a, The MathWorks, Inc., Natick, Massachusetts, USA).

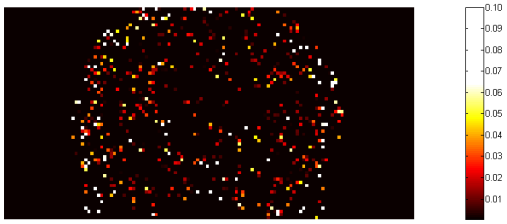
(i) The coronal view

The maximum value of *nox* is $nox_{69} = 877$. Selecting the voxels with $\bar{z}_i > 3.3724$, there are 602 voxels shown in Figure 2.23. The histogram shows the number of outliers in this slice. Thus the 69^{th} slice in coronal view is chosen to analyse. The results of the LLS method are shown in Figure 2.22.

At the voxels with outliers, we recompute the tensor by NLS, LMS and LTS methods and produce the FA map. The difference of FA between estimators is 10^{-2}. In order to see clearly the difference between the two methods, six graphs of FA difference are generated in Figure 2.24.



$|FA_{lls} - FA_{lms}|$



$|FA_{lls} - FA_{lts}|$

Figure 2.22 A DTI image in coronal view.

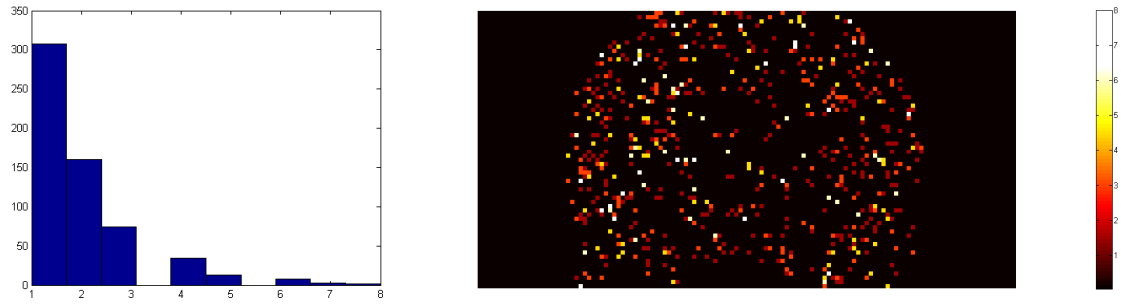
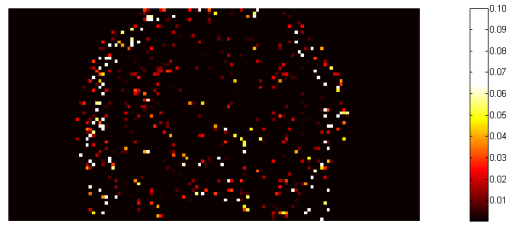


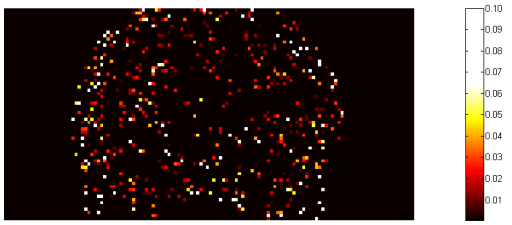
Figure 2.23 Number of outliers at each voxel of Figure 2.22.

g	$\lvert g_{lls} - g_{nls}\rvert$	$\lvert g_{lls} - g_{lms}\rvert$	$\lvert g_{lls} - g_{lts}\rvert$	$\lvert g_{nls} - g_{lms}\rvert$	$\lvert g_{nls} - g_{lts}\rvert$	$\lvert g_{lms} - g_{lts}\rvert$
FA	0.0373 (0.0131)	0.0272 (0.0185)	0.0329 (0.0208)	0.0474 (0.0231)	0.0469 (0.0253)	0.0332 (0.0213)
MD ($\times 10^{-5}$)	7.055 (1.340)	1.128 (0.809)	0.486 (0.002)	8.183 (2.682)	6.569 (1.477)	1.614 (0.704)
ϕ	27.82° (4.05°)	39.74° (8.00°)	42.88° (9.36°)	45.79° (11.47°)	45.13° (11.22°)	41.15° (10.31°)

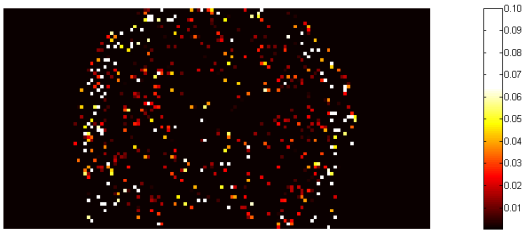
Table 2.2 The mean (median) of absolute value of difference in fractional anisotropy, mean diffusivity and angle of principal fibres among LLS, NLS, LMS and LTS methods on coronal view 69^{th} slice.



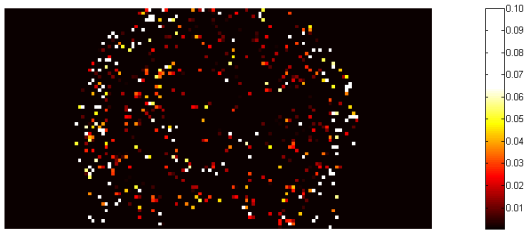
$|FA_{lls} - FA_{nls}|$



$|FA_{lms} - FA_{lts}|$



$|FA_{nls} - FA_{lms}|$



$|FA_{nls} - FA_{lts}|$

Figure 2.24 The difference between each two FA maps.

Table 2.2 summaries the mean and median of the absolute value of difference in their FA, MD and angle of principal directions (v_1). In terms of FA, the difference between all the methods are mostly very small, $0.01 - 0.02$, but there exist some large values, the mean is between 0.02 and 0.04. The largest difference is between the NLS and LMS methods. In MD values, the LLS and LTS have the smallest difference, while the NLS and LMS the largest difference. The LLS has a large difference to NLS estimators, and its result is very close to the LMS and LTS results. When estimating the main direction of fibres, the LLS and NLS are the closest. The traditional methods and robust methods have large differences. The LMS and LTS produce quite different estimators. It is necessary to use both methods. In general, the NLS and LMS methods have the largest difference.

(ii) The sagittal view

Similarly, the maximum number of *noy* is $noy_{65} = 1,334$. Selecting the voxels of $\bar{z}_i > 3.3724$, there are 923 voxels shown in Figure 2.26. We focus on the 65^{th} slice in sagittal view. Figure 2.25 shows the FA map by LLS method.

The results from linear least squares method are the first computed. For the voxels with outliers, we recompute the tensor by non-linear least squares, least median squares and least

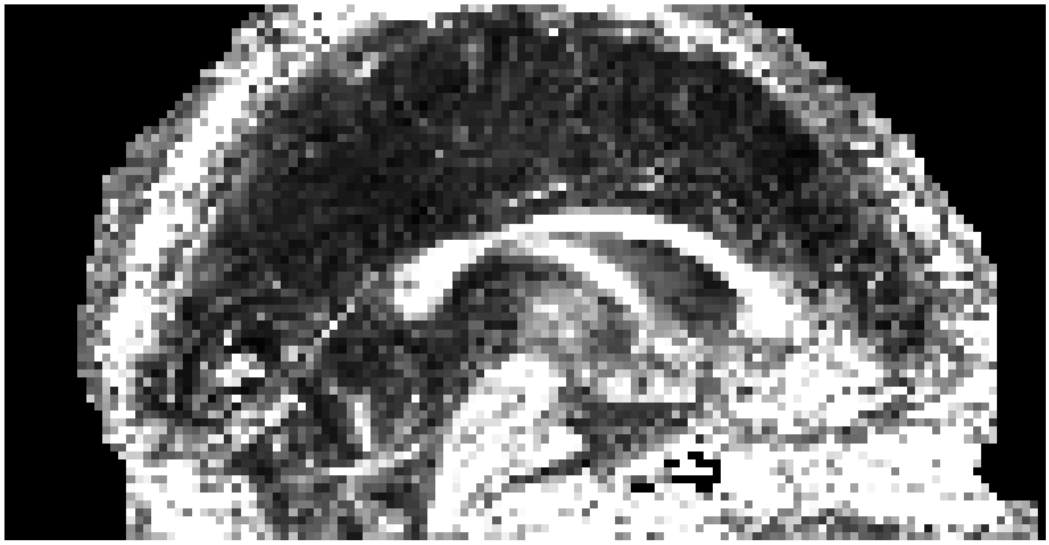


Figure 2.25 A DTI image in sagittal view.

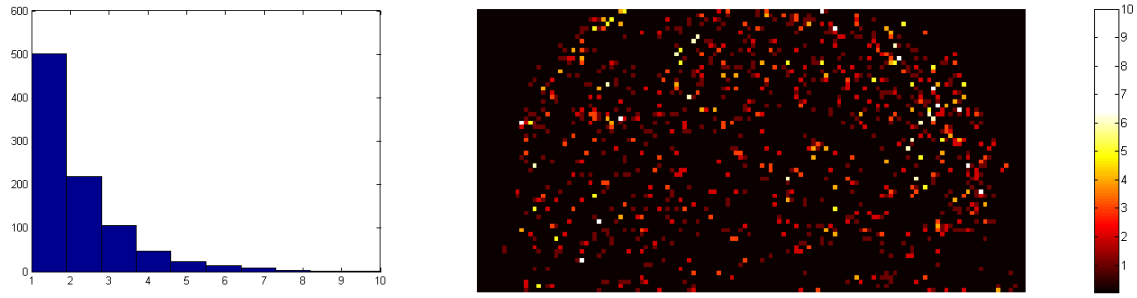
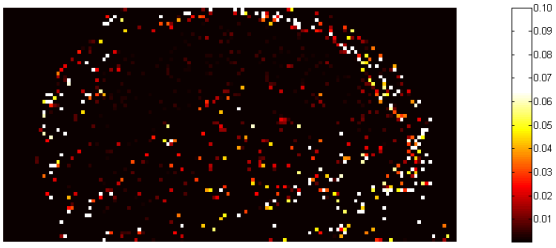


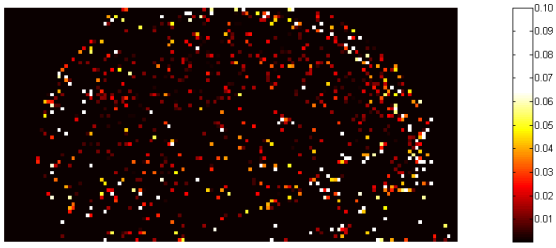
Figure 2.26 Number of outliers at each voxel of Figure 2.25.

g	$\lvert g_{lls} - g_{nls}\rvert$	$\lvert g_{lls} - g_{lms}\rvert$	$\lvert g_{lls} - g_{lts}\rvert$	$\lvert g_{nls} - g_{lms}\rvert$	$\lvert g_{nls} - g_{lts}\rvert$	$\lvert g_{lms} - g_{lts}\rvert$
FA	0.0412 (0.0125)	0.0278 (0.0167)	0.0299 (0.0166)	0.0495 (0.0224)	0.0461 (0.0209)	0.0306 (0.0186)
MD ($\times 10^{-5}$)	6.195 (1.938)	1.767 (0.965)	0.289 (0.204)	7.964 (3.344)	6.484 (1.839)	1.480 (0.668)
ϕ	30.60° (4.86°)	40.78° (9.39°)	45.26° (9.78°)	49.83° (13.80°)	52.69° (13.32°)	44.96° (10.65°)

Table 2.3 The mean (median) of absolute value of difference in fractional anisotropy, mean diffusivity and angle of principal fibres among LLS, NLS, LMS and LTS methods on sagittal view 65^{th} slice.

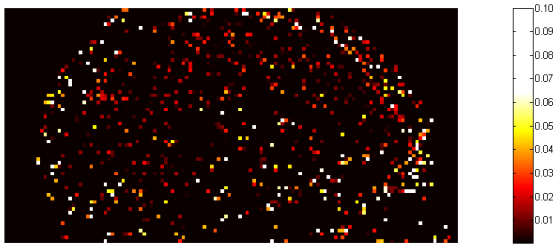


$|FA_{lls} - FA_{nls}|$

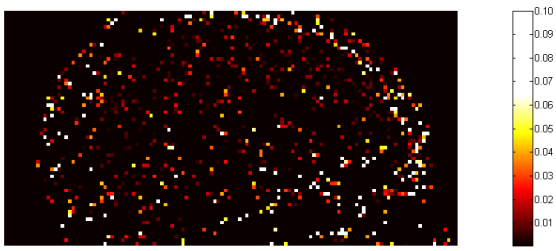


$|FA_{lms} - FA_{lts}|$

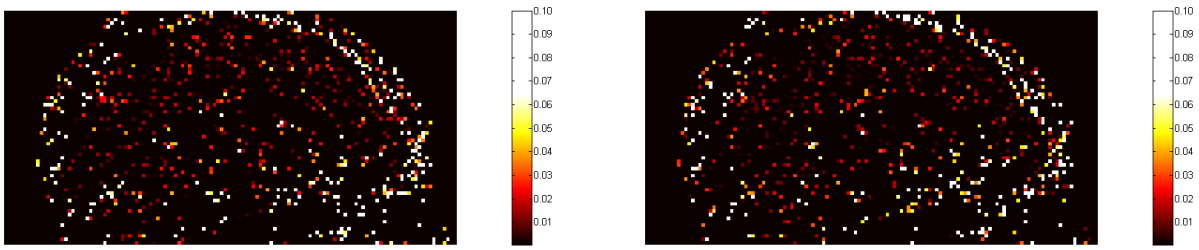
trimmed squares methods. We substitute these results for these corrupted voxels. Figure 2.27 are FA difference between each two methods.



$|FA_{lls} - FA_{lms}|$



$|FA_{lls} - FA_{lts}|$



$|FA_{nls} - FA_{lms}|$ $|FA_{nls} - FA_{lts}|$

Figure 2.27 The difference between each two FA maps.

Table 2.3 summaries the mean and median of the absolute value of difference in terms of their FA, MD and angle of major direction (v_1). The differences of the three properties are left tail skewed. Thus the median is smaller than the mean, especially in FA and angle of major direction. The LLS has a large difference to NLS method than to robust methods in terms of FA and MD. The difference between LLS and robust methods are larger than the difference between LLS and NLS method. The largest differences still come from NLS and LMS method in FA and MD values.

(iii) The transverse view

The maximum value of noz is $noz_{31} = 1,516$. Selecting the voxels with $\bar{z}_i > 3.3724$, there are $1,119$ voxels shown in Figure 2.29. We choose the 31^{th} slice in transverse view to analyse. Figure 2.28 shows the FA map by LLS method.

The results from linear least squares method are the first computed. At the voxels with outliers. We recompute the tensor by non-linear least squares, least median squares and least trimmed squares methods. These corrupted voxels are replaced by the estimators. The difference of FA is illustrated in Figure 2.30.

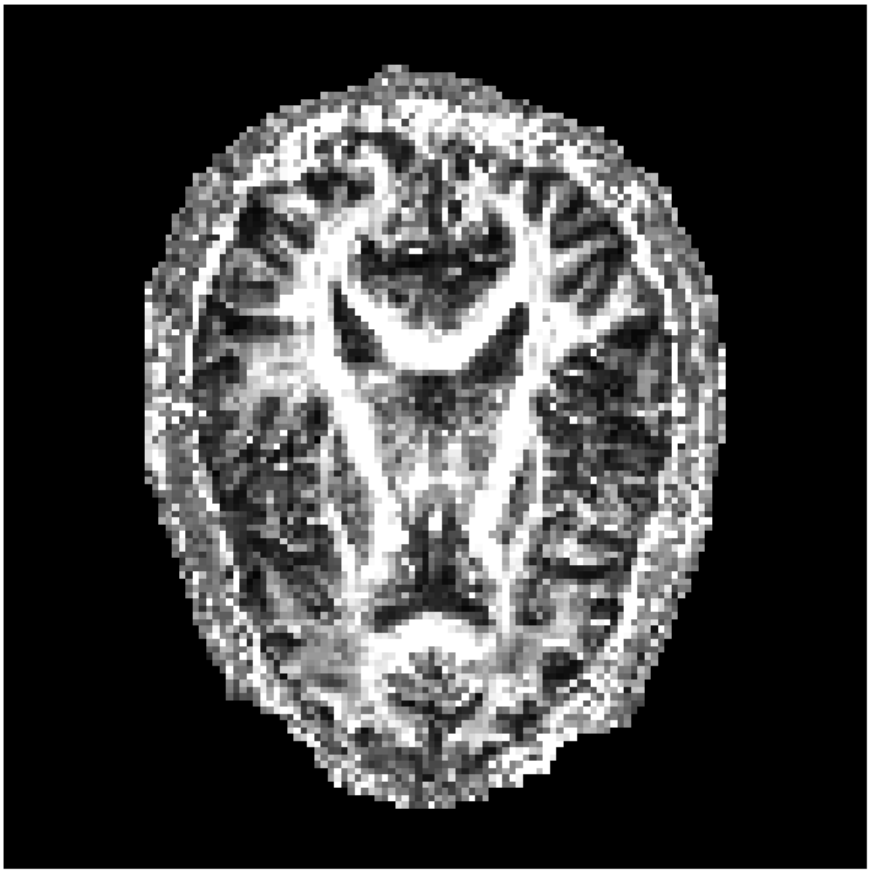


Figure 2.28 A DTI image in transverse view.

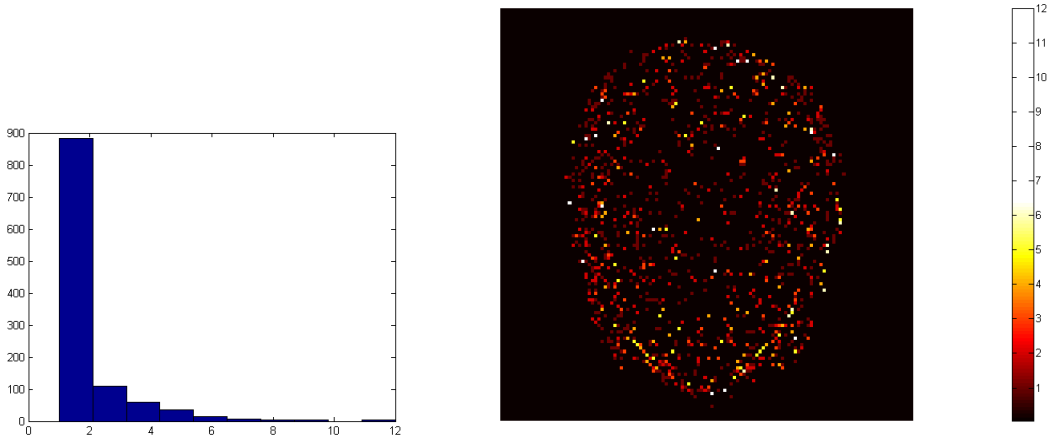
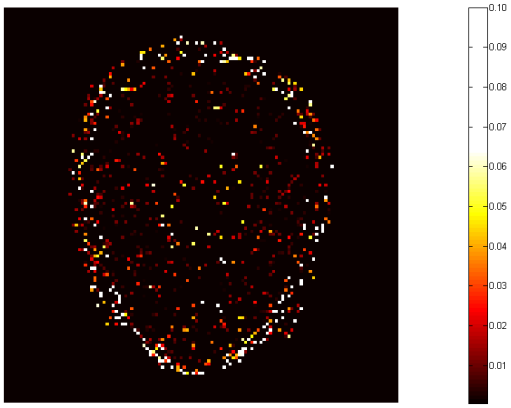


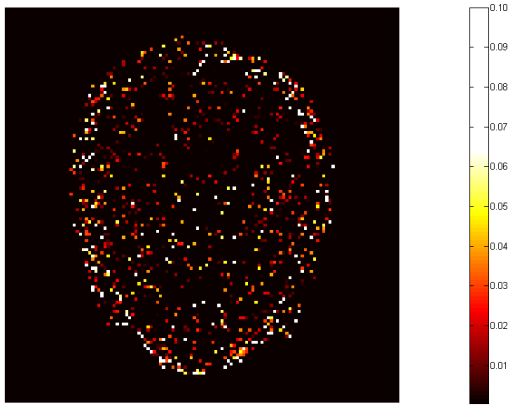
Figure 2.29 Number of outliers at each voxel of Figure 2.28.

g	$\lvert g_{lls} - g_{nls}\rvert$	$\lvert g_{lls} - g_{lms}\rvert$	$\lvert g_{lls} - g_{lts}\rvert$	$\lvert g_{nls} - g_{lms}\rvert$	$\lvert g_{nls} - g_{lts}\rvert$	$\lvert g_{lms} - g_{lts}\rvert$
FA	0.0419 (0.0128)	0.0275 (0.0174)	0.0327 (0.0205)	0.0507 (0.0258)	0.0461 (0.0233)	0.0337 (0.0219)
MD ($\times 10^{-5}$)	6.655 (1.310)	1.017 (0.491)	17.282 (0.177)	7.673 (2.341)	23.937 (1.088)	16.264 (0.621)
ϕ	24.67° (3.63°)	35.83° (7.39°)	39.10° (8.00°)	39.23° (9.90°)	42.24° (9.68°)	37.97° (8.56°)

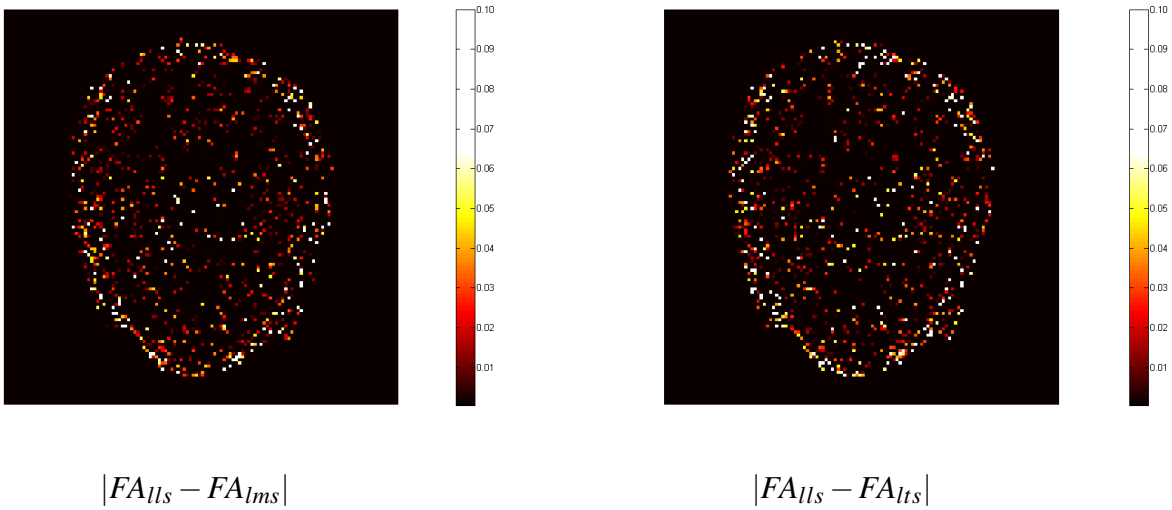
Table 2.4 The mean (median) of absolute value of difference in fractional anisotropy, mean diffusivity and angle of principal fibres among LLS, NLS, LMS and LTS methods on transverse view 31^{th} slice.



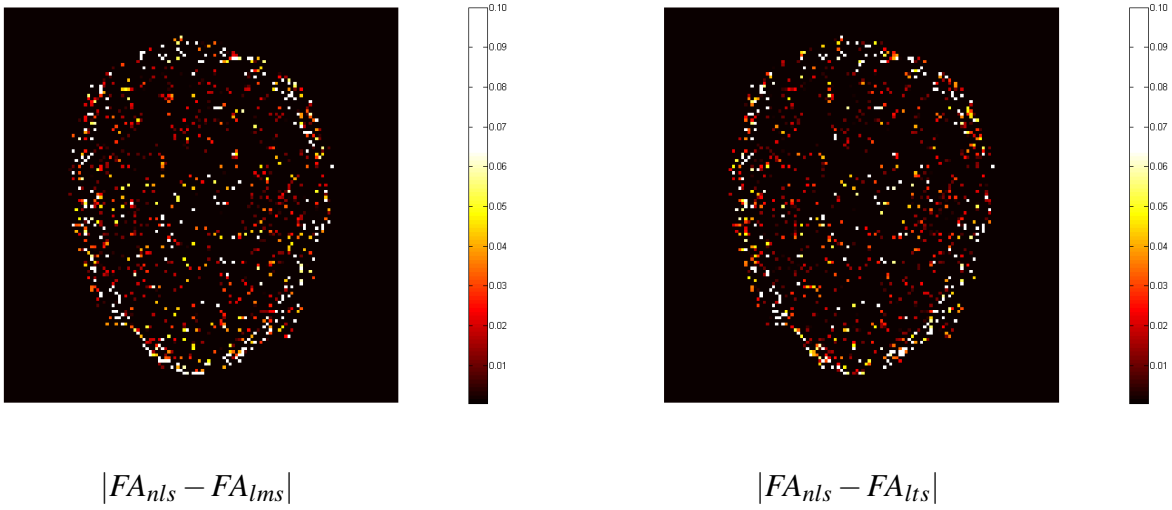
$|FA_{lls} - FA_{nls}|$



$|FA_{lms} - FA_{lts}|$



$|FA_{lls} - FA_{lms}|$ $|FA_{lls} - FA_{lts}|$



$|FA_{nls} - FA_{lms}|$ $|FA_{nls} - FA_{lts}|$

Figure 2.30 The difference between each two FA maps.

Table 2.4 shows the mean (median) of absolute value of difference by four methods on transverse view 31^{th} slice between four methods. The mean of differences is much larger than the median of the difference in terms of FA, MD and main directions. In FA, the NLS and LMS are still the most difference pair. The median of difference between each pair is similar. The MD value from LTS estimator has an extremely difference from other three methods. The difference of main directions of these four methods are similar, the smallest value come from the difference between LLS and NLS methods.

Figure 2.31 The corpus callosum region and its surrounding of a human brain.

2.6.4 Validation

The purpose of this experiment is to validate the four regression methods in the real data. Our ROI is corpus callosum region and its surrounding voxels shown in Figure 2.31. CC is the largest white matter structure in the brain, and it is around 1% of the human brain [112]. We pick a similar size in the study. Our ROI contains 18×22 voxels. The mean of diffusion signal intensities without magnetics is $83,809$ in our DTI ROI data. The mean of signals under magnetics is $15,531$. The standard deviation of noise at the diffusion signals is $2,310.6$. The estimated diffusion tensors are calculated using four regression methods mentioned in this chapter. Then We add Gaussian noise with standard deviation σ to our raw data and estimate the tensors again. The mean of absolute value of differences of properties between the tensors estimated from raw data and noisy data are calculated and the results are shown in Tables 2.5, 2.6, 2.7 and 2.8. These tables are corresponding to four setting levels of noises. The least figures are highlighted in bold in the tables. The regression with the least mean of absolute value of difference indicates that it has the highest robustness.

Table 2.5 shows the mean of absolute value in terms of differences of FA, MD and orientations when $\sigma = 2,000$. The level of noise is approximate to the noise level in the raw data. When the noise is small, LLS method gives the best results, i.e. the smallest mean of absolute value of difference in FA, MD and orientations. The robust methods have slightly greater differences, whilst the NLS method gives the greatest figures.

$\sigma = 2,000$	FA	MD	Orientations
LLS	**0.0521**	**0.0000**	**6.1932**
NLS	0.1302	0.0085	14.7672
LMS	0.0571	0.0001	7.0325
LTS	0.0547	**0.0000**	7.2908

Table 2.5 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 2,000$.

$\sigma = 4,000$	FA	MD	Orientations
LLS	0.1388	0.0002	11.5833
NLS	0.1857	0.0117	18.7787
LMS	0.1331	0.0002	11.8613
LTS	**0.1217**	**0.0001**	**11.1217**

Table 2.6 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 4,000$.

$\sigma = 5,000$	FA	MD	Orientations
LLS	0.1828	0.0003	11.8652
NLS	0.2123	0.0852	19.6930
LMS	0.1730	0.0003	12.8470
LTS	**0.1548**	**0.0002**	**11.4257**

Table 2.7 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 5,000$.

$\sigma = 10,000$	FA	MD	Orientations
LLS	0.4318	0.0009	18.7490
NLS	0.3873	0.0868	23.1357
LMS	0.3326	**0.0007**	18.0376
LTS	**0.3246**	0.0008	**17.0373**

Table 2.8 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated from raw data and noisy data with noise standard deviation $\sigma = 1,0000$.

At the medium level of noise, we have set two σ values. Table 2.6 is under the noise with standard deviation $\sigma = 4,000$ and Table 2.7 shows the figures when $\sigma = 5,000$. Both tables emphasise that the robust methods achieve better results than non-robust methods. The LLS, LMS and LTS methods have similar results, while NLS method has a greater mean of absolute value of difference between the estimated tensors from raw data and noisy data. The LTS method has the least differences of FA, MD and orientations of tensors.

When the level of noise becomes quite large, $\sigma = 10,000$, the robust methods still have slightly better results than non-robust methods in Table 2.8. The LTS method has least differences in terms of FA and orientations and LMS has least difference in terms of MD. On the other hand, the NLS still has the greatest differences in terms of MD and orientations. The LLS have greater mean of absolute value of difference than NLS method in FA.

From this experiment, we conclude that when the noise is small, LLS is the best regression to estimate diffusion tensors. When the noise becomes large, robust methods are better estimations than the non-robust methods in DTI. The LTS is the most robustness among four approaches.

2.7 Summary

In this chapter, we focus on the robust statistics and propose the robust regressions in the estimation of diffusion tensor. The least median squares and least trimmed squares methods with forward search algorithm are analysed in the diffusion tensor imaging Gaussian model to get rid of the effects of unnecessary noise in the measurement scheme. The results are compared with estimators from linear least squares and non-linear least squares methods in simulation studies and applicants. Simulation studies are carried out in three cases: different fractional anisotropy, volume and principal orientations of the tensors. The root-mean-square errors of these properties of estimators are plotted in the figures in presence of outliers counting from 0 up to half of the data. In the real data, we first compare the estimators of a single diffusion tensor and read the pattern of its neighbouring tensors. Then we analyse the real data arising in the diffusion signals and conclude the appropriate outliers selection methods and threshold values. The comparisons are applied to three DTI images from the coronal, sagittal and transverse view respectively. The absolute value of difference in terms of FA, MD and orientations between the various estimators are shown in tables and figures. We also validate the robust and non-robust methods in the real data by adding extra noise to the diffusion signals. The mean of absolute value of differences in FA, MD and orientations are shown between tensors estimated from noisy data and raw data.

In general, the results show that the robust estimators have better results than the traditional estimators. The NLS method is problematic at estimating the determinant. It tends to have a larger volume than the real tensor. The LMS method is preferable when there are more outliers. LTS regression is recommended as it has a better estimation of determinant of tensors than the LMS method. When the noise level is small, the LLS is the best regression to estimate diffusion tensors. When the noise becomes large, robust methods are better estimations than the non-robust methods. The LTS is the most robustness among four methods.

Although the robust estimations are more reliable than non-robust approaches, the traditional statistics is still preferable with usual normal assumptions in absence of derivations from the assumed models. The robust methods are designed to find those outlying points and eliminate or reduce their effects in the range from simple cases with data to detect spotty data in the multivariate studies.

Median computation is time- and space-consuming than traditional regressions. The LTS has a better asymptotic efficiency than LMS and therefore needs even more computation time [90]. Anscombe has pointed out: the robust model insures against accidents caused by deviations from the model based on compromises of some efficiency [4, 38]. Furthermore, there are certain mathematical and numerical difficulties in the LMS and LTS methods. In particular, solutions are not unique and may result in local instabilities [93].

Chapter 3

Bayesian Model with Neighbourhood Priors on Multiple Fibres

The diffusion tensor model is commonly applied under the assumption that the principal eigenvector corresponding to the largest eigenvalue of the tensor, which is aligned along the dominant fibre orientation in a voxel [1]. However, there are regions in the brain where more than one distinct fibre orientations are captured in a single voxel. In DTI studies, how to model the diffusion behaviours at such regions containing more than one distinct fibre orientation is still an open question. Previous researchers have been studied this problem and developed some models which are introduced in Section 3.2. These methods only use the information of the tensor itself and ignore the information of neighbouring tensors. Therefore, we develop Bayesian single and double tensor models which take into account the neighbouring tensors.

3.1 Introduction

Diffusion tensors provide the information of molecular motion under magnetic resonance in the brain. The tensor can be illustrated as an ellipsoid shown in Figure 1.7. The lengths of the axes are proportional to the square root of the eigenvalues and the longest axes of the ellipsoid represent the principal direction of the molecule. The water molecules have a preferable direction along the fibres rather than cross different types of tissues in the brain. The preferable direction commonly indicates the dominant fibre direction. The microstructure and connectivity of fibre orientation can be determined by the motions of the molecules in DTI.

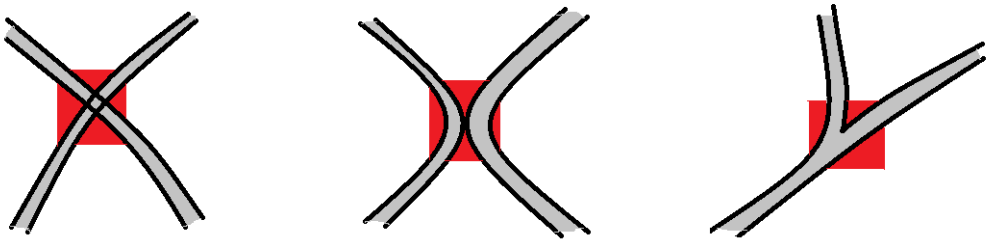


Figure 3.1 Two fibres crossing (right), kissing (middle) and merging (left) of fibre bundles

In DTI, if the fibre bundles orientation has one dominant direction at a voxel, then the water molecule behave along the dominant fibre orientation at such voxel. Recall that the Brownian motion of water molecules follows the Gaussian distribution with zero mean and standard deviation: (Equation (1.5))

$$S_i = S_0 \exp\left(-b g_i^T D g_i\right) + \varepsilon_i \,, \qquad \text{where } i = 1, \cdots, N.$$

In the Gaussian model, it is assumed that each voxel contains a single, coherently oriented fibre and it is not precisely described the multi-tensor orientation situation. It describes a single tensor model. In the brain, there are regions of more than one distinct fibre orientation in a voxel. It has been estimated that more than 90% of white matter voxels contains crossing fibres [51]. The fibre bundles may be crossing, kissing or merging (divergence and convergence), which is shown in Figure 3.1. At such voxels, the one tensor model is no longer satisfied. It is important to show the connectivity in the white matter more precisely. Considering its constraints in the intersections of fibres, the multi-fibre identifiability problem is the focus of this chapter.

Previous researchers have different approaches at the multiple fibre voxels, we will review their methods in the next section.

DTI is a powerful tool to analysis the tissues in the human brain. However, such a framework (Equation (1.5)) is that each voxel contains a single, coherently oriented fibre and it does not precisely describe the multi-tensor orientation situations.

This chapter will discuss the double tensor models using neighbour's information, especially those one tensor voxels neighbourhoods, and propose Bayesian Model using Neighbour's Information (BN) Framework. Bayesian single model, Bayesian double model and model selection are deduced in the methodology. We then define what are the neighbours of

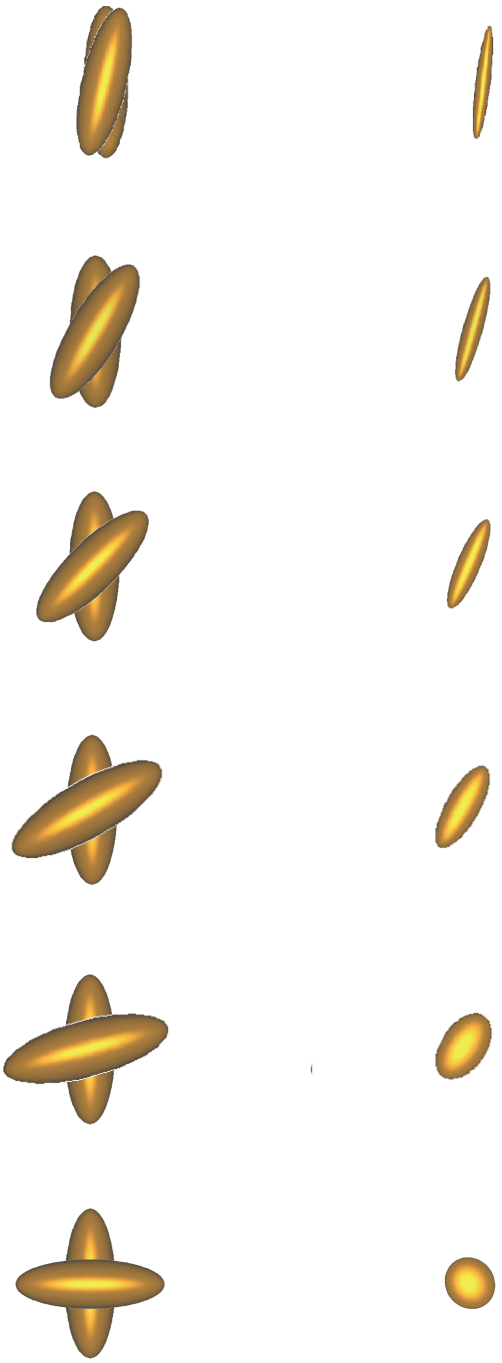


Figure 3.2 The same tensors crossing at 15°, 30°, 45°, 60°, 75° and 90°, and their estimators in single tensor model

a voxel, and the three categories of voxels: corner voxels, boundary voxels and inner voxels. The starting point and phase portrait are discussed and the Bayesian neighbour algorithm will be illustrated in the diagram. The simulation studies are performed to examine the dependence of the angle between two tensors, mean diffusivity, fractional anisotropy and test the goodness of fit. A real DTI image from a healthy human brain in coronal view is presented. The region of pons and corpus callosum is enlarged and the Bayesian models and BN algorithm are performed in the real data.

3.2 Multiple tensors models

At crossing fibres, a variety of reconstruction algorithms has been considered by previous researchers. These methods include the high angular resolution diffusion weighted imaging (HARDI), diffusion orientation distribution function (ODF), mixture Gaussian models, Bayesian model and spherical deconvolution approaches.

3.2.1 High angular resolution diffusion weighted imaging

Many methods have been proposed based on the idea of extracting more information from the diffusion weighted imaging signals (DWI). One of the most common tools is the high angular resolution diffusion weighted imaging (HARDI) acquisition protocol. HARDI captures the higher angular frequency features of the diffusive weighted signals, and it is not adequately modelled by a single diffusion tensor [106]. It gathers a much larger number of uniformly distributed diffusion weighted gradient directions than DTI.

3.2.2 Mixture Gaussian

In a crossing fibre region, the mixture Gaussian model assumes that the signal measured is the sum of the diffusion weighting signals from each distinct fibre bundle. The probability density function p of diffusion process of a particle displacements x over a fixed time t is assumed as a zero-mean trivariate Gaussian distribution [3]

$$G(x;D,t) = [(4\pi t)^3 det(D)]^{-\frac{1}{2}} \exp(-\frac{x^T D^{-1} x}{4t}) \tag{3.1}$$

where **D** is the diffusion tensor. A more generalised form for p is summation of Gaussian densities with weights $a_i \in [0,1]$ and $\sum_i a_i = 1$,

$$p(x) = \sum_{i=1}^{n} G(x; D_i, t) \tag{3.2}$$

Thus the fourier transform of $p(x)$ of the signals is

$$A(q) = \sum_{i=1}^{n} a_i \exp(-tq^T D_i q) \tag{3.3}$$

3.2.3 Diffusion spectrum imaging (DSI)

The main focus of diffusion spectrum imaging is to construct a representation p from signal measurements via a fast Fourier transform straightforward. The orientation distribution function (ODF) [94] is

$$B(\hat{x}) = \int_0^\infty p(\alpha\hat{x}) d\alpha \tag{3.4}$$

where $\hat{x}$ is a unit vector in the direction of displacements x. Such ODF is the radial projection of p onto the unit sphere. When the p has the greatest value, the function B has the greatest value in the direction and DSI assumes thus direction is the fibre direction. The function can have multiple pairs of equal peaks but in different directions, which indicates the crossing fibres cases.

The results show that such function has its peaks at the expected crossing fibres regions [110]. Thus it has an advantage on solving multi-fibre problems in diffusion tensor imaging. However the DSI requires sufficient details in magnitude measurements, it is a disadvantage that the acquisition times are longer. The results also show that the function also has peaked in unexpected grey matter areas. This may be caused by the measurement noise or anatomic structure.

3.2.4 q-Ball Imaging

The q-ball imaging method proposed by Tuch [105] estimates the orientation distribution function by Funk transformation of diffusion signal at a fixed radius in q-space. The value of the Funk transform of a spherical function at the point $\hat{x}$ is the integral of the function over

the great circle $C(\hat{x})$ perpendicular to $\hat{x}$. The q-ball approximates

$$B(\hat{x}) = \int_{C(\hat{x})} A(Q\hat{q})d\hat{q} \tag{3.5}$$

where Q is the mean or some typical $|q_i|$. The signal at each $\hat{x}$ and its radial basis function are

$$B(\hat{x}) = \sum_{k=1}^{K} \beta_k \theta_k(\hat{x})\theta_k(\hat{x}) = \exp[-(Dcos^{-1}(|\hat{x}\hat{y}_k|))^2] \tag{3.6}$$

where **D** is a constant scaling parameter and $\hat{y}_k$ is the unit vector evenly distributed on the unit sphere y_k. The equations can be solved by transforming to a matrix form [105]. In the q-ball imaging, there is a Fourier relationship between the spin propagator (the spin displacement probability density function) and the distribution of the diffusion weighted signal over q-space. Since it needs to measure the diffusion weighted signal over 3D q-space (as in diffusion spectrum imaging (DSI)), it leads to long scan times.

3.2.5 Spherical deconvolution methods

Another approach is the unified deconvolution framework proposed by Jian and Vemuri [52]. It assumes that the diffusion signal is the convolution of the fibre orientation distribution f. It is a distribution rather than a discrete fibre orientations. Let $\hat{x}$ be the orientation, then f is

$$f(\hat{x}) = \sum_{i=1}^{N} \beta_i \theta_i(\hat{x}). \tag{3.7}$$

The signal

$$A(q) = \sum_{i=1}^{N} (\beta_i \int A_f(q,\hat{x})\theta_i(\hat{x})d\hat{x}). \tag{3.8}$$

We can compute the tensors using linear methods. The equation can be rewritten as $A = XB$, where $A = [A(q_1), \cdots, A(q_N)]^T$ is the vector of normalised signals, $B = (\beta_1, \cdots, \beta_i)^T$ is the vector of functions weights, and X is the matrix with $X_i = \int A_f(q,x)\theta_i(\hat{x})d\hat{x}$.

3.3 Methodology

In the previous section, we have reviewed the recently proposed techniques for the multiple tensor problems. All of these methods estimate a double tensor based on the diffusion signals of the voxel itself. The behaviour of neighbouring tensors has not been considered. The fibre bundles are connected in the brain. The orientation of two fibre bundles implies the two major directions of the double tensors voxels at the crossing of these two fibre bundles. The behaviours of neighbouring tensors, for example, the fractional anisotropy, mean diffusivity, the principal orientation and etc, are vital references to the estimated tensors. The properties of the estimated double tensors and the surrounding tensors should be coincident or similar. Hence we proposed a Bayesian method that takes the neighbouring tensors into account to estimate the tensors of the multiple fibres. In this section, the background of Bayesian statistics will be first introduced, then the Bayesian single tensor model and Bayesian double tensor model will be proposed. Finally, Occam's razor and model selection of Bayesian single or double tensor model will be described.

3.3.1 Bayesian statistics

The foundation of Bayesian statistics is Baye Theorem. It states that the conditional probability of A given by B is proportional to the product of its own probability and the probability of B given by A.

$$P(A|B) \propto P(B)P(B|A) \tag{3.9}$$

Suppose some unknown quantities $\theta = (\theta_1, \theta_2, \cdots, \theta_n)$, where n is a positive integer. A priori belief about θ is expressed in terms of the pdf $p(\theta)$. Now suppose there is some observations data relevant to their values, call it $X = (X_1, X_2, \cdots, X_n)$, which have a probability distribution that depends on these n unknown quantities as parameters. Thus the pdf of the vector X depends on the vector θ, i.e. $p(X|\theta)$.

Considering both your piror beliefs about θ and the observed data, based on the Bays' Theorem, we have

$$p(\theta|X) \propto p(\theta)p(X|\theta), \tag{3.10}$$

thus $p(X|\theta)$ is a density function of X for fixed θ. We call $p(X|\theta)$ the *likelihood function*, written as

$$l(\theta|X) = p(X|\theta) \tag{3.11}$$

It is a function of θ. Sometimes the *log-likelihood* function is used for estimating,

$$\mathscr{L}(\theta|X) = \log l(\theta|X) \tag{3.12}$$

The Bayes' theorem indicates that the posterior probability density function for parameters is proportional to the density for parameters with prior belief multiple by the likelihood for parameters given the data [63].

$$posterior\ density \propto likelihood\ function \times prior\ density \tag{3.13}$$

Note that it is possible that the prior beliefs about θ differs, and therefore the posterior belief also differs. However, all the prior beliefs should be bounded in the form of $p(X|\theta)$, i.e. related to the data θ. When reasonable large data have been collected, the posterior beliefs will usually be very close.

The Bayesian methods can be applied sequentially. For example, the initial observations samples X of parameter θ,

$$p(\theta|X) \propto p(\theta)l(\theta|X) \tag{3.14}$$

now there is a second set of observations Y distributed independently of the first sample. then

$$p(\theta|X,Y) \propto p(\theta)l(\theta|X,Y) \tag{3.15}$$

The independence implies that

$$p(X,Y|\theta) \propto p(X|\theta)p(Y|\theta) \tag{3.16}$$

and

$$l(\theta|X,Y) \propto l(\theta|X)l(\theta|Y) \tag{3.17}$$

Therefore the posterior for θ given X and Y by treating your posterior given X as the prior for the observation Y,

$$p(\theta|X,Y) \propto p(\theta)l(\theta|X)l(\theta|Y) \propto p(\theta|X)l(\theta|Y) \tag{3.18}$$

Edwards, Lindman and Savage mentioned Bayesian probability is that "probability is orderly opinion, and that inference from data is nothing other than the revision of such opinion in the light of relevant new information" [27]. Indeed, comparing to statisticians

consider the probability of a hypothesis in restricted circumstances, the Bayesian statisticians prefer to consider the probability as a hypothesis, then modify these prior beliefs under the relevant data which have observed to achieve posterior beliefs [72, 63].

3.3.2 Single tensor model

Recall the single tensor Gaussian model (Equation (1.5)), it assumes the each voxel contains a single, coherently oriented fibre. It describes a single tensor model. The fibre bundles orientation has one dominant direction at each voxel, and the water molecule follows the Gaussian distribution with zero mean and standard deviation along the dominant fibre orientation.

The previous researchers have collected prior information from the tensors then updated the information by signal intensities. The posterior distribution is calculated by specified likelihood functions. In our research, we consider the problem based on two facts:

- The single neighbour has known fibre orientations.
- The fibre orientations of their neighbours should be the same or very similar.

The Bayesian framework parameterises the diffusion tensor as the product a 3×3 matrix and its transpose. Such parameters are estimated by Bayesian model with suitable prior distributions. To ensure the symmetric positive semi-definiteness of tensor **D**, we reparameterise **D** as a product of a general 3×3 matrix and its transpose, $D = B^T B$. For example, the cholesky decomposition $B = chol(D)$, or square root matrix $B = D^{\frac{1}{2}}$.

$$S_i = S_0 \exp\left(-b g_i^T B^T B g_i\right) + \varepsilon_i \,, \qquad \text{where } i = 1, \cdots, N. \tag{3.19}$$

where the Gaussian noise $\varepsilon_i \sim \mathcal{N}(0, \sigma_1^2)$. Subsequently, the diffusion signals S_i are also Gaussian variables with mean $S_0 \exp\left(-b g_i^T B^T B g_i\right)$ and variance σ_1^2. The parameter space for the single tensor model contains two variables, the parameter B and the variance σ_1^2.

First consider the variable B, which is also a 3×3 matrix with nine unknown entries.

$$B = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix} \tag{3.20}$$

We define the vectorised $\vec{B}$ as a column vector by stacking the columns of B,

$$\vec{B} = \begin{pmatrix} B_{11} & B_{21} & B_{31} & B_{12} & B_{22} & B_{32} & B_{13} & B_{23} & B_{33} \end{pmatrix}^T \tag{3.21}$$

Assuming the prior distribution of B is a multivariate Gaussian distribution according to our initial beliefs, the tensor at the current voxel should follow a similar pattern of one of its neighbours,

$$\vec{B} \sim \mathcal{N}_9(B_N, \xi^2 I_9)\sigma_1^2 \sim Inv-Gamma(\alpha, \beta) \tag{3.22}$$

where $\vec{B}_1$ is the neighbour tensor of B, I_9 is the identity matrix. The probability distribution function of B is as follows:

$$P(B) = (2\pi\xi^2)^{-\frac{9}{2}} \exp\left\{-\frac{1}{2\xi^2}(\vec{B}-\vec{B}_N)^T(\vec{B}-\vec{B}_N)\right\} \tag{3.23}$$

If ξ takes large value, the prior B is in high uncertainty.

Suppose the variance σ^2 and the parameter B are independent, a suitable prior for Gaussian noise σ_1^2 is [Lee]

$$\sigma_1^2 \sim Inv\text{-}Gamma(\alpha, \beta) \tag{3.24}$$

Specifically, the distribution is

$$P(\sigma_1^2) = \frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma_1^2)^{-\alpha-1}\exp(-\frac{\beta}{\sigma_1^2}) \tag{3.25}$$

Thus the prior distribution of single tensor model is the joint distribution of B and σ^2, i.e. the product of their distribution as they are independent

$$P(B, \sigma_1^2) = (2\pi\xi^2)^{-\frac{9}{2}}\frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma_1^2)^{-\alpha-1}\exp\left\{-\frac{\beta}{\sigma_1^2} - \frac{1}{2\xi^2}(\vec{B}-\vec{B}_N)^T(\vec{B}-\vec{B}_N)\right\} \tag{3.26}$$

The prior distribution represents a function of our unknown parameters, and the likelihood function collects the data from our observations, the measured DTI image data. For the single tensor model, the likelihood function of B and σ^2 by given DTI signal intensities S is

$$l(B, \sigma^2|S) = \prod_{i=1}^{N} f(S_i|B, \sigma^2) \tag{3.27}$$

$$= (\frac{1}{\sqrt{2\pi\sigma_1^2}})^N \exp\left\{-\frac{1}{2\sigma_1^2}RSS_1\right\} \tag{3.28}$$

where N is the number of observations and RSS_1 is the residuals sum of squares between the observed data and the expected signals.

$$RSS_1 = \sum_{i=1}^{N} [S_i - S_0 \exp(-b g_i^T B^T B g_i)]^2 \tag{3.29}$$

The posterior distribution $P(B, \sigma_1^2 | S)$ is obtained by Bayes' theorem,

$$P(B, \sigma_1^2 | S) = \frac{(2\pi)^{-\frac{N+9}{2}} \beta^\alpha}{\Gamma(\alpha)} \xi^{-9} (\sigma_1^2)^{-\frac{N+2\alpha+2}{2}} \exp\left\{ -\frac{2\beta + RSS_1}{2\sigma_1^2} - \frac{(\vec{B} - \vec{B}_N)^T (\vec{B} - \vec{B}_N)}{2\xi^2} \right\} \tag{3.30}$$

B and σ^2 are acquired by maximising the posterior distribution. Alternatively, it can be estimated by sampling the posterior distribution by Markov chain Monte Carlo simulation. It is more nature to maximising the log-posterior distribution,

$$\log P(B, \sigma_1^2 | S) = -\frac{N + 2\alpha + 2}{2} \log(\sigma_1^2) + \left\{ -\frac{2\beta + RSS_1}{2\sigma_1^2} - \frac{(\vec{B} - \vec{B}_N)^T (\vec{B} - \vec{B}_N)}{2\xi^2} \right\} \tag{3.31}$$

Thus the diffusion tensor $\mathbf{D}$ is given by $D = B^T B$.

3.3.3 Double tensor model

There is more than one distinct fibre orientation in some regions of the brain. The multiple tensor model should be considered. As most of the crossing, kissing and merging regions contains two distinct fibres, we only finalised the double tensor cases. Such model of tensors $D_1 = B_1^T B_1$ and $D_2 = B_2^T B_2$ is expressed as

$$S_i = S_0 \exp(-b g_i^T B_1^T B_1 g_i) + S_0 \exp(-b g_i^T B_2^T B_2 g_i) + \varepsilon_i, \quad \text{where } i = 1, \cdots, N. \tag{3.32}$$

where ε_i is the Gaussian noise follows by $\mathcal{N}(0, \sigma_2^2)$, thus the signal intensities S_i also follows Gaussian distribution with variance σ_2^2.

In a double tensor voxel, either two distinct single fibre orientations come across at this voxel, or follows the same pattern of a nearby double fibre orientations. Assuming we have two neighbours D_{N1} and D_{N2}, which can be expressed as $D_{N1} = B_{N1}^T B_{N1}$ and $D_{N2} = B_{N2}^T B_{N2}$. The prior belief of our vectorised $\vec{B}_1$ and $\vec{B}_2$ follows the Gaussian distribution with mean $\vec{B}_{N1}$ and $\vec{B}_{N2}$ respectively, and with the same variance of $\xi_1^2 I_9$,

$$\vec{B}_1 \sim \mathscr{N}_9(\vec{B}_{N1}, \xi_1^2 I_9) \tag{3.33}$$

$$\vec{B}_2 \sim \mathscr{N}_9(\vec{B}_{N2}, \xi_1^2 I_9) \tag{3.34}$$

The probability density functions of $\vec{B}_1$ and $\vec{B}_2$ are specifically

$$P(B_1) = (2\pi\xi_1^2)^{-\frac{9}{2}} \exp\left\{ -\frac{1}{2\xi_1^2}(\vec{B}_1 - \vec{B}_{N1})^T(\vec{B}_1 - \vec{B}_{N1}) \right\} \tag{3.35}$$

and

$$P(B_2) = (2\pi\xi_1^2)^{-\frac{9}{2}} \exp\left\{ -\frac{1}{2\xi_1^2}(\vec{B}_2 - \vec{B}_{N2})^T(\vec{B}_2 - \vec{B}_{N2}) \right\} \tag{3.36}$$

The joint density distribution of B_1 and B_2 is a multivariate Gaussian distribution of 18 unknown parameters,

$$\mathscr{J}_{B_1,B_2} \sim \mathscr{N}_{18}(\mu_{B_1,B_2}, \Sigma) \tag{3.37}$$

where the mean μ_{B_1,B_2} is a column vector with stacking the elements of $\vec{B}_{N1}$ and $\vec{B}_{N2}$,

$$\mu_{B_1,B_2} = \begin{pmatrix} \vec{B}_{N1} \\ \vec{B}_{N2} \end{pmatrix} \tag{3.38}$$

and the covariance Σ is the 18×18 matrix partitioned as follows,

$$\Sigma = \begin{pmatrix} \mathrm{var}(B_1) & \mathrm{cov}(B_1, B_2) \\ \mathrm{cov}(B_1, B_2) & \mathrm{var}(B_2) \end{pmatrix} \tag{3.39}$$

As B_1 and B_2 are not necessarily independent, we need to explicitly specificate the $\mathrm{cov}(B_1, B_2)$. Assuming that

$$\vec{B}_1 - \vec{B}_2 \sim N_9(\vec{B}_{N1} - \vec{B}_{N2},\ \xi_2^2 I_9) \tag{3.40}$$

The variance of $(\vec{B}_1 - \vec{B}_2)$ can be expressed by its definition

$$\begin{aligned}
\text{var}[\vec{B}_1 - \vec{B}_2] &= E[(B_1 - B_2)^2] - (E[B_1 - B_2])^2 \\
&= E[B_1^2 + B_2^2 - 2B_1B_2] - (\vec{B}_{N1} - \vec{B}_{N2})^2 \\
&= E[B_1^2] + E[B_2^2] - 2E[B_1B_2] - (\vec{B}_{N1} - \vec{B}_{N2})^2 \\
&= \text{var}(B_1) + (E[B_1])^2 + \text{var}(B_2) + (E[B_2])^2 - 2E[B_1B_2] \\
&\quad -(\vec{B}_{N1} - \vec{B}_{N2})^2 \\
&= \xi_1^2 I_9 + \vec{B}_{N1}^T \vec{B}_{N1} + \xi_1^2 I_9 + \vec{B}_{N2}^T \vec{B}_{N2} - 2E[B_1B_2] \\
&\quad -(\vec{B}_{N1}^T \vec{B}_{N1} - 2\vec{B}_{N1}^T \vec{B}_{N2} + \vec{B}_{N2}^T \vec{B}_{N2}) \\
&= 2\xi_1^2 I_9 + 2\vec{B}_{N1}^T \vec{B}_{N2} - 2E[B_1B_2]
\end{aligned}$$

From Equation (3.40), the variance of $(\vec{B}_1 - \vec{B}_2)$ is assumed as

$$\text{var}[\vec{B}_1 - \vec{B}_2] = \xi_2^2 I_9 \tag{3.41}$$

Combining these two equations of $\text{var}[\vec{B}_1 - \vec{B}_2]$, we obtain the expectation of the product of two variables,

$$E[B_1B_2] = (\xi_1^2 - \frac{1}{2}\xi_2^2) I_9 + \vec{B}_{N1}^T \vec{B}_{N2} \tag{3.42}$$

Consequently, the covariance of B_1 and B_2 is as follows

$$\begin{aligned}
\text{cov}(B_1, B_2) &= E[(B_1 - E[B_1])(B_2 - E[B_2])] \\
&= E[B_1B_2] - E[B_1]E[B_2] \\
&= (\xi_1^2 - \frac{1}{2}\xi_2^2) I_9 + \vec{B}_{N1}^T \vec{B}_{N2} - \vec{B}_{N1}^T \vec{B}_{N2} \\
&= (\xi_1^2 - \frac{1}{2}\xi_2^2) I_9
\end{aligned}$$

The covariance of B_1 and B_2 are only related to our assumption of their variance. The choice of the neighbour tensors $\vec{B}_{N1}$ and $\vec{B}_{N2}$ does not matter.

Now we can state the variance of the joint distribution $\mathscr{J}_{B_1,B_2}$ is

$$\Sigma = \begin{pmatrix} \xi_1^2 I_9 & (\xi_1^2 - \frac{1}{2}\xi_2^2) I_9 \\ (\xi_1^2 - \frac{1}{2}\xi_2^2) I_9 & \xi_1^2 I_9 \end{pmatrix} \tag{3.43}$$

The joint distribution of B_1 and B_2 is verified as

$$P(B_1, B_2) = (2\pi\xi_1^2)^{-9} \exp\left\{ -\frac{1}{2\xi_1^2} (\vec{B}_1 - \vec{B}_{N1})^T \Sigma^{-1} (\vec{B}_2 - \vec{B}_{N2}) \right\} \tag{3.44}$$

where

$$\Sigma^{-1} = \frac{1}{\xi_2^2(\xi_1^2 - \frac{1}{4}\xi_2^2)} \begin{pmatrix} \xi_1^2 I_9 & -(\xi_1^2 - \frac{1}{2}\xi_2^2) I_9 \\ -(\xi_1^2 - \frac{1}{2}\xi_2^2) I_9 & \xi_1^2 I_9 \end{pmatrix} \tag{3.45}$$

Denote the exponential part in Equation (3.44) as Θ, thus

$$\begin{aligned}
\Theta &= \frac{1}{2} (\vec{B}_1 - \vec{B}_{N1})^T \Sigma^{-1} (\vec{B}_1 - \vec{B}_{N1}) \\
&= \frac{1}{2\xi_2^2(\xi_1^2 - \frac{1}{4}\xi_2^2)} \Big\{ \xi_1^2 (\vec{B}_1 - \vec{B}_{N1})^T (\vec{B}_1 - \vec{B}_{N1}) I_9 \\
&\quad -(\xi_1^2 - \frac{1}{2}\xi_2^2)(\vec{B}_2 - \vec{B}_{N2})^T (\vec{B}_1 - \vec{B}_{N1}) I_9 + (\xi_1^2 - \frac{1}{2}\xi_2^2)(\vec{B}_1 - \vec{B}_{N1})^T (\vec{B}_2 - \vec{B}_{N2}) I_9 \\
&\quad +\xi_1^2 (\vec{B}_2 - \vec{B}_{N2})^T (\vec{B}_2 - \vec{B}_{N2}) I_9 \Big\} \\
&= \frac{1}{2\xi_2^2(\xi_1^2 - \frac{1}{4}\xi_2^2)} \Big\{ \xi_1^2 \Big[(\vec{B}_1 - \vec{B}_{N1}) + (\vec{B}_2 - \vec{B}_{N2}) \Big]^T \Big[(\vec{B}_1 - \vec{B}_{N1}) + (\vec{B}_2 - \vec{B}_{N2}) \Big] \\
&\quad +\xi_2^2 (\vec{B}_1 - \vec{B}_{N1})^T (\vec{B}_2 - \vec{B}_{N2}) \Big\} \\
&= \frac{1}{\xi_2^2(1+\rho)} \Big[(\vec{B}_1 - \vec{B}_{N1}) + (\vec{B}_2 - \vec{B}_{N2}) \Big]^T \Big[(\vec{B}_1 - \vec{B}_{N1}) + (\vec{B}_2 - \vec{B}_{N2}) \Big] \\
&\quad +\frac{1}{\xi_1^2(1+\rho)} (\vec{B}_1 - \vec{B}_{N1})^T (\vec{B}_2 - \vec{B}_{N2})
\end{aligned}$$

where ρ is the correlation of coefficient of $\vec{B}_1$ and $\vec{B}_2$,

$$\rho = \frac{\mathrm{cov}(\vec{B}_1, \vec{B}_2)}{\sigma_{\vec{B}_1} \sigma_{\vec{B}_2}} = \frac{\xi_1^2 - \frac{1}{2}\xi_2^2}{\xi_1^2} = 1 - \frac{\xi_2^2}{2\xi_1^2}. \tag{3.46}$$

Since $-1 \leq \rho \leq 1$, we have $0 \leq \xi_2^2 \leq 4\xi_1^2$, and $\vec{B}_1$ and $\vec{B}_2$ is independent only if $\xi_2^2 = 2\xi_1^2$.

The prior belief of variance, σ_2^2, is assumed to follow a inv-Gamma distribution,

$$\sigma_2^2 \sim Inv - Gamma(\alpha, \beta) \tag{3.47}$$

The likelihood function of priors B_1,B_2,σ_2^2 given by their signals S_i is as follows:

$$l(B_1, B_2, \sigma_2^2|S) = (\frac{1}{\sqrt{2\pi\sigma_2^2}})^N \exp\left\{-\frac{1}{2\sigma_2^2} RSS_2\right\} \tag{3.48}$$

where RSS_2 is the residuals sum of squares between the observed signal intensities and the expected signals.

$$RSS_2 = \sum_{i=1}^{N} [S_i - S_0 \exp(-b g_i^T B_1^T B_1 g_i) - S_0 \exp(-b g_i^T B_2^T B_2 g_i)]^2 \tag{3.49}$$

Consequently, the posterior distribution of double tensor model using neighbours tensors as priors is

$$p(B_1, B_2, \sigma_2^2|S) = \frac{(2\pi)^{-\frac{N}{2}-9}\beta^\alpha}{\Gamma(\alpha)} (\xi_1^2)^{-9} (\sigma_2^2)^{-\frac{N}{2}-\alpha-1} \exp\left\{-\frac{2\beta + RSS_2}{2\sigma_2^2} - \frac{\Theta}{2\xi_1^2}\right\} \tag{3.50}$$

Alternative, the log-posterior distribution of double tensor model is

$$\log p(B_1, B_2, \sigma_2^2|S) = -\frac{N + 2\alpha + 2}{2} \log(\sigma_2^2) + \left\{-\frac{2\beta + RSS_2}{2\sigma_2^2} - \frac{\Theta}{2\xi_1^2}\right\} \tag{3.51}$$

The parameter B_1 and B_2 can be obtained by maximising the posterior or log-posterior functions, and the double tensors D_1 and D_2 are given by $D_1 = B_1^T B_1$ and $D_2 = B_2^T B_2$.

3.3.4 Occam's razor and model selection

After calculating the double tensor voxels using neighbour's information in a region, it is important to justify the number of fibre orientations in each voxel. It is normally considered that there are one or two fibre orientations at each voxel in DTI. We need to decide whether to use the single tensor model or double tensor model to simulate each tensor. Thus the number of single neighbours and double neighbours of voxels can be specified in Φ.

Occam's razor refers to *the hypotheses with the fewest assumptions should be selected*.The fewer adjustable parameters for a hypothesis will result in an enhanced posterior probability [48]. In Bayesian statistics, the prior is not required for a simpler model to avoid bias [70]. The Bayesian Occam's Razor is testing the likelihood hypotheses rather than the priors.

One the straight-forward method for model selection in Bayesian statistics based on the Bayes factor. can be used for model selection. The single tensor model (M_1) is given as

$$S_i = S_0 \exp(-bg^T QQ^T g_i) + \varepsilon_i \tag{3.52}$$

and double tensor model (M_2) is

$$S_i = S_0 \exp(-bg^T Q_1 Q_1^T g_i) + S_0 \exp(-bg^T Q_2 Q_2^T g_i) + \varepsilon_i \tag{3.53}$$

and the Bayes factor K in terms of M_1 and M_2 is the maximised likelihood of M_2 divided by the maximised likelihood of M_1 [12]

$$K(S) = \frac{L(M_2|S)}{L(M_1|S)} = \frac{p(M_2|S)p(M_1)}{p(M_1|S)p(M_2)} \tag{3.54}$$

The Bayes factor shows a preference between these two model. If K_B is greater than one, it suggests the stronger selection of M_2 over M_1 by the data. The Jeffrey's scale shows that the thresholds of 3 are the optimal choice for the Bayesian methods in Table 3.1 [49].

Bayes factor K	Strength of evidence
$K < 10^0$	Supports for M_1
$10^0 < K < 10^{1/2}$	Very weak evidence for M_2
$10^{1/2} < K < 10^1$	Substantial evidence for M_2
$10^1 < K < 10^{3/2}$	Strong evidence for M_2
$10^{3/2} < K < 10^2$	Very strong evidence for M_2
$K_B > 10^2$	Decisive for M_2

Table 3.1 Jeffrey's scale of evidence for Bayes factors.

Another alternative table for interpretation of K is Kass and Raffery's scale.

$2\ln K$	Bayes factor$K$	Strength of evidence
0 to 2	$1 < K < 3$	Very weak evidence for M_2
2 to 6	$3 < K < 20$	Positive evidence for M_2
6 to 10	$20 < K < 150$	Strong evidence for M_2
> 10	$K > 150$	Very strong evidence for M_2

Table 3.2 Kass and Raffery's scale of evidence for Bayes factors.

Bayesian information criterion (BIC) or Schwarz criterion is another factor model selection. [G. Schwarz]

$$\mathrm{BIC}_i = -2\ln L(M_i|S) + k\ln(n) \tag{3.55}$$

where $L(M_i|S)$ is the maximised likelihood for model M_i, n is the number of observations and k is the number of free parameters to be estimated. We prefer the model with the lowest BIC. BIC is only valid when the sample size is much larger than the number of parameters in the model, i.e. $n > k$. It has a preference variable for low-dimension problems rather than high dimensions.

Another similar criterion is the Akaike information criterion (AIC). The model with the minimum AIC value is preferable. AIC only shows a relative value of goodness of fit and the complexity of the model.

$$\mathrm{AIC}_i = 2k - 2\ln L(M_i|S) \tag{3.56}$$

Note before estimating a diffusion tensor in a voxel, it is unknown whether it is a single tensor model or a double tensor model. The criterion of model selection is used to judge such information.

3.4 Algorithm

Estimating the tensor field with single and multiple fibres, the results will be influenced if calculating the tensors approaching from different phase portraits. In this section, we also develop the Bayesian neighbours (BN) algorithm for our Bayesian tensor model estimation using neighbour's information as priors. First, we will specify the neighbours of a voxel, then discuss the starting point and phase portrait, the BN algorithm and the procedure of the framework.

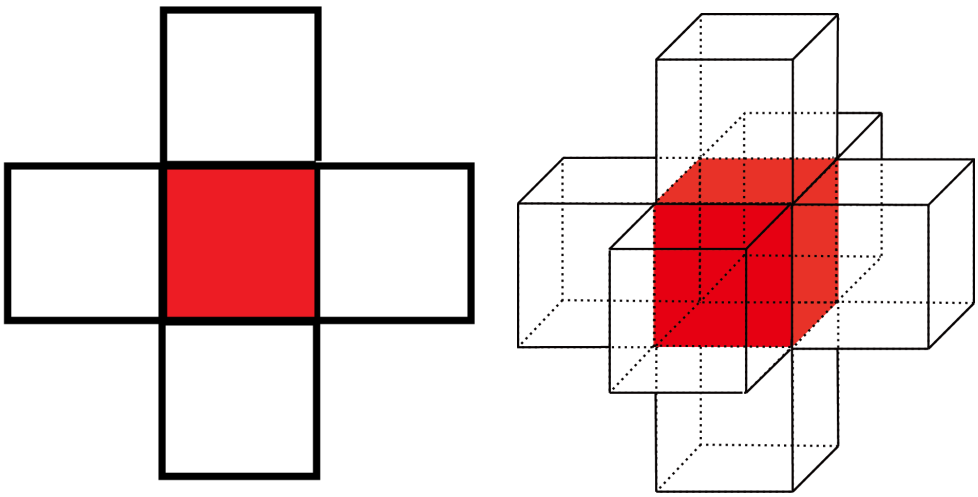


Figure 3.3 The neighbours of a voxel in 2D and 3D.

3.4.1 Neighbours and voxels

The main focus in our framework is to use neighbours' information to estimate a multiple tensor in a voxel. It is necessary to specify the neighbours of a voxel. In our algorithm, we define a *neighbour* of a voxel is adjoined to the voxel and share the same boundary or side. In 2D, a voxel haa four neighbours, including up, down, left, right. There are six neighbours of a voxel in 3D, including up, down, left, right, forward and backwards. Figure 3.3 describes these two situations in graphs. If a neighbour of a voxel is a single tensor model, defined as a *single neighbour*. If a neighbour is a double tensor model, it is called a *double neighbour*.

Suppose we are analysing a region which contains voxels have multiple fibre orientations. Let the multiple tensor model voxels in such region form a space Φ. In Φ, we define voxels which have one single neighbour as *boundary voxels*. The *corner voxels* have two or more single neighbours. Note both the corner voxels and boundary voxels have single neighbours, they are *edge voxels*. The voxels without any single neighbours are called *inner voxels*. Figure 3.4 illustrates these three categories of voxels in a 8×8 cube and one of its interior slice. The corner voxels are at the edges of the cube in yellow. The boundary voxels are at the central of surfaces of the cube in white. The inner voxels are invisible in the cube but they can be seen from the interior slice of the cube coated in blue. Table 3.3 summaries three categories of voxels and the number of their single voxels.

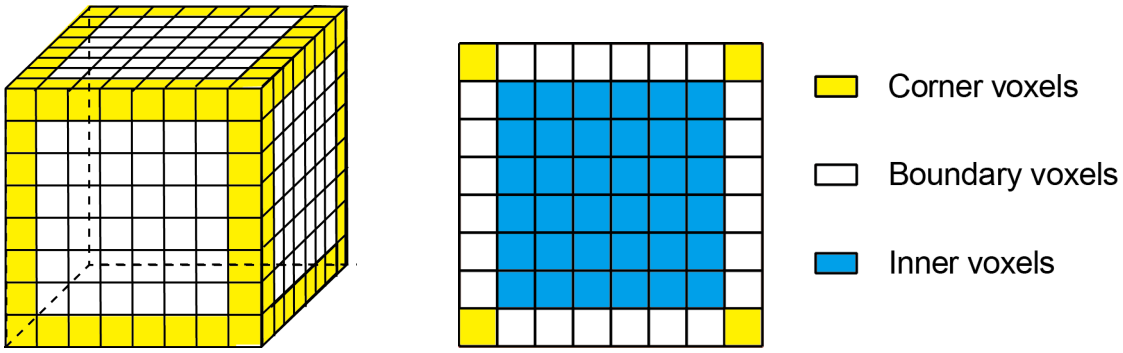


Figure 3.4 The three categories of voxels in 3D and its interior slice

Voxels	Number of single neighbours
Corner voxel	2 or 3
Boundary voxels	1
Inner voxels	0

Table 3.3 Summaries of three categories of voxels and the number of theirs single neighbours.

3.4.2 Starting point and phase portrait

Our objection is to estimate the voxels in Φ using the signals in these voxels as the likelihood and their neighbours' information as priors. The starting point of the procedure should be the edge voxels. They ensure their neighbour voxels containing at least one or more single tensor models. Assuming such regions are a square or cube in Figure 3.3. In reality, the double tensor model regions are not necessary a standardised cube. They may be any shape. We try to extract more information from the neighbours, thus any edges voxel with at least two distinct fibre orientation is better considered as an initial voxel, i.e. edge voxels.

We start from the outermost layer, which is the edge voxel. After estimating all the voxels in this layers. These voxels contain known estimated tensors. The next layer will be the outermost layer in the inner voxels region. In other words, the outermost layer for the unestimated tensors. We repeat the procedure until we reach the centre of the estimated regions.

Each voxel will be fitted with the single tensor model with the four alternatives of priors: tensor of a single neighbour, two tensors of a double neighbour, the median of all neighbours and a random tensor. The best estimation will be determined by the largest corresponding

Bayes factor. If the Bayes factor exceeds the preset threshold value, the voxel will be considered as a single model voxel. Otherwise, the voxel will be fitted with the double tensor model. The pair of priors can be two tensors of a double neighbour or two random tensors. In addition, we also collect all single or double tensors neighbours and classify into two groups by their orientations. The median covariance tensor will be obtained in each group. Thus we have another pair of priors in the double tensor model. The Bayes factor is also the judge of the best double tensor estimations. Finally, the single tensor estimation and the double tensor estimation will be compared by Bayes factor, BIC or AIC. The best model will be selected for the voxel. If the model selection indicates for a single tensor model, the tensors remain the same. If the model selection criterion indicated for a double tensor model, the voxel information will be altered to double tensors.

3.4.3 Bayesian neighbour's algorithm

The Bayesian model using neighbour's information (BN) Algorithm is used to process the tensor using Bayesian double tensor models tensors and its neighbouring information as prior beliefs. It is stated below:

1. Start from the unestimated outermost voxels;
2. Collect the neighbours of each voxel;
3. Use single tensor model fitting with priors: tensor of a single neighbour, two tensors of a double neighbour, the median of all neighbours and a random tensor;
4. Choose the best single model estimation using test criterion;
5. If the likelihood of the single model exceeds the threshold value, determination of single model, jump to step 9, otherwise, continue the next step;
6. Use double tensor model fitting with priors: two tensors of a double neighbour, median of group orientations of all neighbours or two random tensors;
7. Choose the best double model estimation using test criterion;
8. Determination of the single tensor model or double tensor model, update with estimated double tensors for double voxels and update with estimated single tensors if the **D** is not SPD for single voxels;
9. After estimating all the voxels at the outermost layer, back to step 1.

The procedure of the BN framework is also illustrated in the Figure 3.5

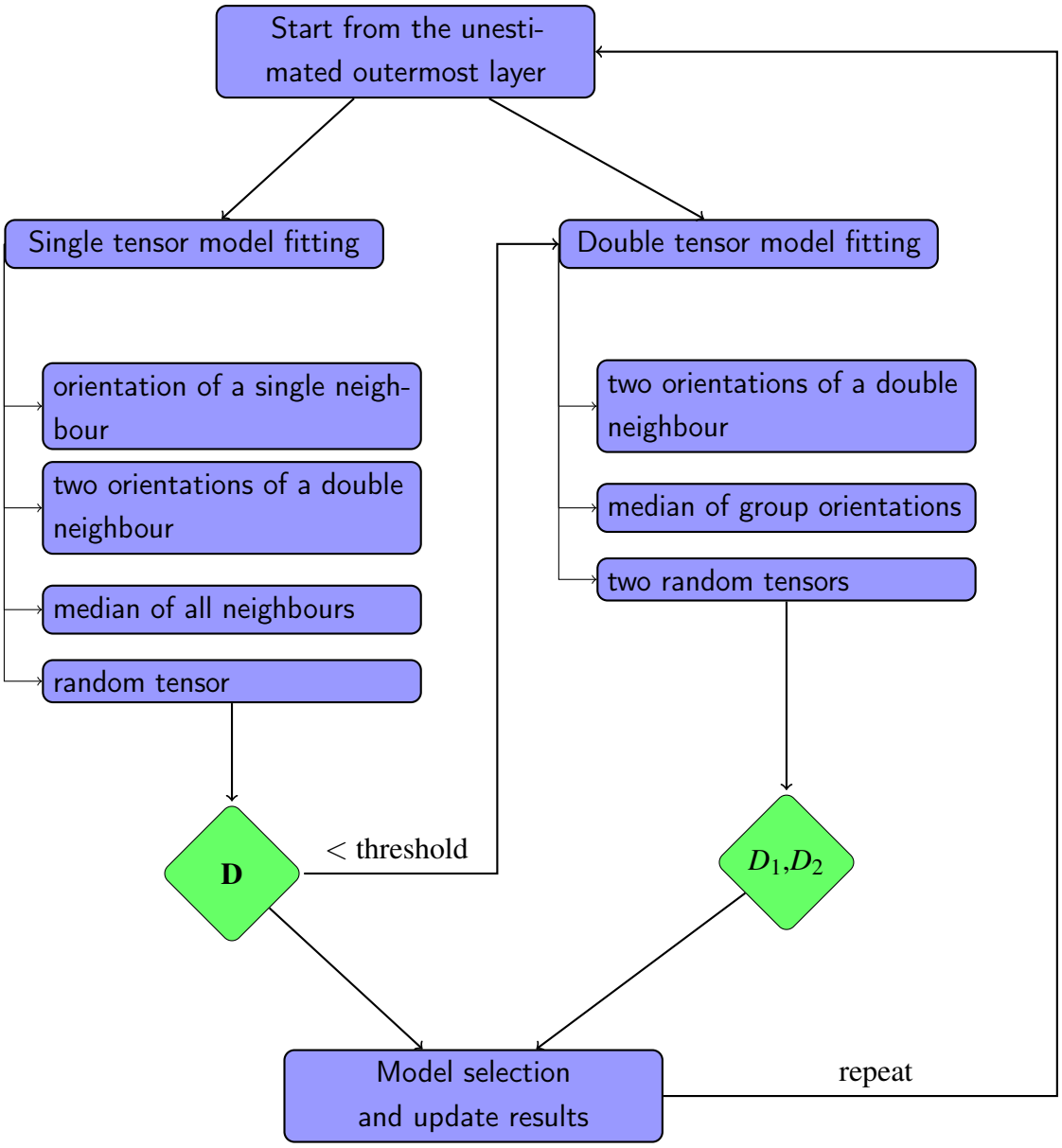


Figure 3.5 The diagram of BN algorithm framework.

3.5 Simulation studies

We perform the BN algorithm to obtain figures and graphs for visualisation using MATLAB(R2013a, The MathWorks, Inc., Natick, Massachusetts, USA). The function *fminsearch* is used to solve the maximum posterior optimisation problem. In this section, we simulated tensors fields that consist of single and double tensors at the crossing regions. Then the estimations of tensors will be analysed in terms of fractional anisotropy, mean diffusivity and principal orientations. Four simulation studies are performed. The first three simulation studies are to examine the dependence of the estimation on the FA and MD of two simulated tensors and the angle between the two principal diffusion orientations respectively. The last simulation study is focused on measuring the goodness of fit of the estimated double tensors.

3.5.1 Dependence of angle between two tensors

The three simulation studies in Sections 3.5.1, 3.5.2 and 3.5.3, are to examine the dependence of the estimation on the properties of tensors in this section. They are designed to examine the dependence of FA, MD and the angle of two principal orientations respectively. In these three simulation studies, we fix one double tensor D_1, and consider three settings of another diffusion tensor, D_{2a}, D_{2b} and D_{2c}. Consider a 5×5 sample size simulating the tensors at the crossing regions of two distinct fibre orientations. The outermost tensors are single tensor voxels, while the central 3×3 are double tensors voxels. The tensors obey the same direction at each fibre. A Gaussian noise of 0.5 is adding on the Cholesky decomposition of diffusion tensors D_1 and D_2 in the simulated sample tensors, we form diffusion tensors B_1 and B_2. The tensor field is formed by 3×3 double voxels with B_1 and B_2 at the centre and their 5×5 single neighbours. Simulate the diffusion signals $S_i, i = 1, \cdots, 32$ based on the tensor field with $S_0 = 10,000$ and $\varepsilon = 100$.

We measure the dependence of a property of diffusion tensors and fix the other two properties among FA, MD and angle each time. The BN algorithm is applied on the simulated tensor field to capture the diffusion behaviour at each voxel. The Bayesian single model will be used to simulated the single neighbours, and simulated tensors are the initials of the Bayesian double models when estimating the double tensors. B_1 and B_2 are simulated true tensors, BN_1 and BN_2 are estimated Bayesian double tensors. The mean, median and standard deviation of FA, MD and angle with D_1 and D_2 are compared respectively.

In the first simulation study, we measure the dependence of angle between two distinct fibre orientations. The study is carried out between two distinct fibres orientations crossing

at 30°, 60° and 90°. We fix the diffusion tensors with eigenvalues are 1,3 and 1, and the corresponding eigenvectors are along x, y and z axis respectively. The FA value measuring the diffusion anisotropy is 0.6030 and the MD measuring the size of the diffusion tensor is 1.6667. The defined double tensors are as follows:

$$
\begin{aligned}
D_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
D_{2a} &= \begin{pmatrix} 1.5 & 0.866 & 0 \\ 0.866 & 2.5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
D_{2b} &= \begin{pmatrix} 2.5 & 0.866 & 0 \\ 0.866 & 1.5 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
D_{2c} &= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

Figure 3.6 shows the simulated tensors D_1, D_{2a}, D_{2b} and D_{2c}. The results of three estimated tensor fields are shown in Figure 3.7. The estimated double tensors are best illustrated when the two tensors are crossing at 90°. The shape of one tensor in the double voxels tends to be slightly smaller or larger when the crossing angle is 60°. When the angle of the tensors gets smaller, to 30°, the volume (i.e. determinant) of some tensors tends to be much larger than its true value.

The scatter plots of FA, MD and the angle between the simulated tensors and estimated tensors between two distinct fibre orientations crossing at 30°, 60° and 90° are shown in Figure 3.11. The scatter plots of FA, MD and angle in Figure 3.11(a) is more diverse, while it is more clustering in Figure 3.11(c). The FA tends to be larger than its true value at 30°. At 60°, the scatter plot of FA values between the estimated tensors and the simulated tensors performs a linear relationship. At 90°, the variation of the FA is quite small. In terms of MD, the value tends to be larger in three cases. The smaller the angle between the fibre

Figure 3.6 The simulated tensors D_1, D_{2a}, D_{2b} and D_{2c} in Section 3.5.1.

Figure 3.7 The estimated tensors field in Section 3.5.1.

orientations, the larger the estimated MD. Basically, the scatter plots of the angle in three different angles are illustrated in the equal relationship.

Tables 3.4, 3.5 and 3.6 give the absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors of these three angles. The standard deviation of the FA, MD and angle are much larger at 30° than 60° and 90°. Comparing the tensors crossing at 90° and 60°, the mean and standard deviation of the difference of FA and MD in $B_1 - D_1$ becomes smaller while the difference of FA and MD in $B_2 - D_2$ becomes larger.

Overall, when two fibres are crossing at the relatively larger angle, the Bayesian double tensor model works quite well. The determinants and the mean diffusivity tends to be larger when the difference of the two fibre orientations are small, and it is tricky to determine whether there is one fibre or two fibres.

	$\lvert B_1 - D_1 \rvert$			$\lvert B_2 - D_{2a} \rvert$		
	mean	median	σ	mean	median	σ
FA	0.0406	0.0293	0.1176	0.0217	0.0095	0.1160
MD	0.0806	0.0536	0.8189	0.4650	0.1205	1.0269
$\phi(*, D_1)$	7.5910	7.9141	4.0930	8.8389	6.8067	4.6811

Table 3.4 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between estimated tensors and their neighbours if they are crossing at 30°.

	$\lvert B_1 - D_1 \rvert$			$\lvert B_2 - D_{2b} \rvert$		
	mean	median	σ	mean	median	σ
FA	0.0154	0.0266	0.0569	0.0067	0.0135	0.0333
MD	0.0319	0.0033	0.2738	0.0711	0.0059	0.2573
$\phi(*, D_1)$	5.4759	4.9851	1.4649	5.8522	5.7282	3.5026

Table 3.5 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between estimated tensors and their neighbours if they are crossing at 60°.

	$\lvert B_1 - D_1 \rvert$			$\lvert B_2 - D_{2c} \rvert$		
	mean	median	σ	mean	median	σ
FA	0.0294	0.0362	0.0145	0.0033	0.0034	0.0367
MD	0.0423	0.0008	0.2234	0.0262	0.0768	0.2117
$\phi(*, D_1)$	2.5568	2.4149	1.8076	2.4662	2.0471	1.6902

Table 3.6 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and their neighbours if the two distinct fibres are crossing at 90°.

3.5.2 Dependence of mean diffusivity

The second simulation study is to measure the dependence of MD of diffusion tensors. The study is carried out between two distinct fibres orientations crossing at 90°. We fix the diffusion tensors with eigenvalues 1, 3 and 1, and the corresponding eigenvectors are along x, y and z axis respectively. The FA value measuring the diffusion anisotropy is 0.6030 and the MD measuring the size of the diffusion tensor is 1.6667. The second tensors have the same

FA value and with MD value: 0.1667, 0.5556 and 1.6667. The ratio of MD value between the D_1 and D_{2*} is 10, 3 and 1 respectively. The defined double tensors are as follows:

$$
\begin{aligned}
D_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
D_{2a} &= \begin{pmatrix} 0.3 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{pmatrix} \\
D_{2b} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.3333 & 0 \\ 0 & 0 & 0.3333 \end{pmatrix} \\
D_{2c} &= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

The scatter plots of FA, MD and angle between simulated tensors and estimated tensors are shown in Figure 3.11. When two tensors have similar MD values, the estimators perform better. When the MD of one tensor is three times of another tensor, the MD of the tensor which has larger MD tends to be much larger. When the MD of one tensor is ten times of another tensor, the deviation of FA and MD are large, and the results are not desirable.

3.5.3 Dependence of fractional anisotropy

In the previous two simulations we mainly considered the situations between two linear tensors. In the next simulation studies, we focus on the dependence of FA. The experiment is between three settings of FA value in diffusion tensor D_{2*}. We consider three extreme cases: linear tensors (the eigenvalues $\lambda_1 >> \lambda_2 \approx \lambda_3$), planar tensors ($\lambda_1 \approx \lambda_2 >> \lambda_3$) and spherical tensors ($\lambda_1 \approx \lambda_2 \approx \lambda_3$). The study is carried out between two distinct fibres orientations crossing at 90°. We fix the diffusion tensors with eigenvalues of 2, 3 and 1, and the corresponding eigenvectors are along x, y and z axis respectively. The FA value

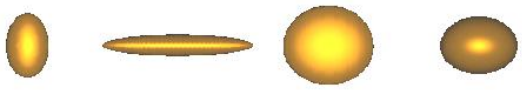


Figure 3.8 The simulated tensors D_1, D_{2a}, D_{2b} and D_{2c}.

measuring the diffusion anisotropy is 0.4629 and the MD measuring the size of the diffusion tensor D_1 is 2. The second tensors, D_{2a}, D_{2b} and D_{2c}, have the same MD value and with FA value: 0.8402, 0.5053 and 0.1715. The defined double tensors are as follow:

$$
\begin{aligned}
D_1 &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
D_{2a} &= \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
D_{2b} &= \begin{pmatrix} 4.2 & 0 & 0 \\ 0 & 3.7 & 0 \\ 0 & 0 & 1.1 \end{pmatrix} \\
D_{2c} &= \begin{pmatrix} 3.6 & 0 & 0 \\ 0 & 2.7 & 0 \\ 0 & 0 & 2.7 \end{pmatrix}
\end{aligned}
$$

Figure 3.8 shows the simulated tensors D_1, D_{2a}, D_{2b} and D_{2c}. The scatter plots of FA, MD and angle between simulated tensors and estimated tensors if the two tensors are in

linear, planar or spherical shapes are shown in Figure 3.11. The results are optimal when two tensors are linear tensors. The FA, MD and angle of the estimated tensors are very close to those of simulated tensors. When one of the tensor is of the planar shape, the MD tends to be larger and the angle tends to be smaller. When a tensor is isotropic, i.e. in the spherical shape, the FA and the MD of the estimated tensors have larger results. Considering the spherical and planar tensors in the brain, they indicate that the water molecule does not have a preferential orientation. The Bayesian framework is more apter under the assumption of anisotropy tensors.

3.5.4 Goodness of fit

The simulation study in this section is focused on measuring the goodness of fit of the estimated double tensors. The size of the simulated tensor field is corresponding to our region of interest (ROI) in the brain. The largest white matter structure in the brain is corpus callosum (CC), around 1% of the human brain [112]. Thus a similar size data is simulated in our experiments. The same tensors D_1, D_{2a}, D_{2b} and D_{2c} are considered at crossing 30°, 60° and 90°. The 3D cases follow the same patterns of the 2D data. Two slices of tensor fields are added before and behind of the corresponding three crossing angles, so the 2D data will be the central slice in the 3D data. There are $18 \times 18 \times 18$ voxels and the central $16 \times 16 \times 16$ are double voxels. The Bayesian double tensors B_1 and B_2 are Gaussian noise setting with zero mean, $sd = 0.1$ on eigenvalues, $sd = 4^\circ$ on angles. The estimated tensor field will be compared with the simulated tensor field.

Figure 3.9 shows the estimated tensor field of two fibres crossing at 30°, 60° and 90°. The results are similar to the 2D case. The determinant and MD of some tensors tends to be larger than their true values when the two fibre orientations are crossing at 30°. The estimated double tensors are similar to their simulated tensors when the angle is 60° and 90°.

The results of absolute value of differences in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors of these three angles are compared in Table 3.7, 3.8 and 3.9. The mean, median and standard deviation of the difference in FA, MD and angle of two fibre orientations get larger when the angle of two tensors gets smaller. Notice Table 3.5 contains slightly conflicting results that the difference between FA and MD of B_1 and D_1 is smaller than the results in Table 3.6. That is because the data set is small. When the data set is large enough, we will obtain the same conclusion.

	$\lvert B_1 - D_1 \rvert$			$\lvert B_2 - D_{2a} \rvert$		
	mean	median	σ	mean	median	σ
FA	0.0512	0.0500	0.1543	0.0423	0.0132	0.1451
MD	0.3962	0.0176	1.2673	0.4730	0.107e	0.9905
$\phi(*, D_1)$	8.9045	10.2671	5.1623	8.2428	8.0634	4.4297

Table 3.7 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors.

	$\lvert B_1 - D_1 \rvert$			$\lvert B_2 - D_{2b} \rvert$		
	mean	median	σ	mean	median	σ
FA	0.0192	0.0179	0.0623	0.0078	0.0065	0.0381
MD	0.1748	0.1905	0.4582	0.0498	0.1391	0.3009
$\phi(*, D_1)$	5.3921	4.1099	2.1562	4.9131	3.9311	4.021

Table 3.8 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors.

	$\lvert B_1 - D_1 \rvert$			$\lvert B_2 - D_{2c} \rvert$		
	mean	median	σ	mean	median	σ
FA	0.0102	0.0098	0.0277	0.0079	0.0066	0.0386
MD	0.0723	0.0523	0.2011	0.0031	0.0467	0.1942
$\phi(*, D_1)$	3.2511	3.1199	2.0191	3.0021	2.7980	2.3182

Table 3.9 The absolute value of difference in terms of fractional anisotropy, mean diffusivity and angle of principal fibres between simulated tensors and estimated tensors.

3.6 Real data analysis

A DTI image from a healthy human brain in coronal view and a partial enlarged ROI are shown in the Figure 3.9. The enlarged ROI delineated pons and corpus callosum. With FA map as background, the principal orientations of tensors are shown by red lines at each voxel in the ROI. The BN algorithm is performed on the ROI to estimate the diffusion tensors. The information of the neighbouring voxels are used as prior believes for single and double tensor model. The Bayes factor of threshold value 3 is used for model selection and the FA of tensor is also considered as a reference. The fibre orientations of each tensor are shown in red line, coded with the same background in Figure 3.10. The transverse pontine fibres can be seen with left-right orientation whin the pons. The corticospinal tracts are with inferior-superior orientation. The graph shows the intersection of these two fibres.

There are $31 \times 13 = 403$ voxels in the selected enlarged ROI in Figure 3.9. There are 302 double voxels, i.e. 74.94% of the whole voxels. We also add Gaussian noise with standard deviation $\sigma = 2,000$ to diffusion signals at the enlarged ROI. The BN algorithm is performed again to estimate the diffusion tensors of signals with extra noise. By the determination of Bayes factor, there are 290 double voxels from noisy data. The double voxel is around 71.96% of the whole voxels, 2.98% less than the result from raw data. As the percentage drops quite small, we consider the BN algorithm is robust.

3.7 Summary

In Chapter 3, we focus on the intersection of the multiple fibres in human brain. We have reviewed the recent techniques of multiple tensor models problems, including high angular resolution diffusion weighted imaging, orientation distribution function, mixture Gaussian model, diffusion spectrum imaging, q-Ball imaging and spherical deconvolution methods. As the information of neighbouring tensors is potentially important in relation to the estimated multiple tensors in the intersection and has been ignored in the previous mentioned methods, we develop the Bayesian single and double tensor models taken into the neighbouring tensors. The background of Bayesian statistics has been introduced. In the Bayesian models, the priors uses the information of neighbouring tensors, and the variance is represented by the inverse Gamma distribution. The likelihood function and posterior distribution of single and double tensor models have been deduced. Several indices of model selection criterion are stated. We also proposed Bayesian neighbour algorithm to estimate the single and double

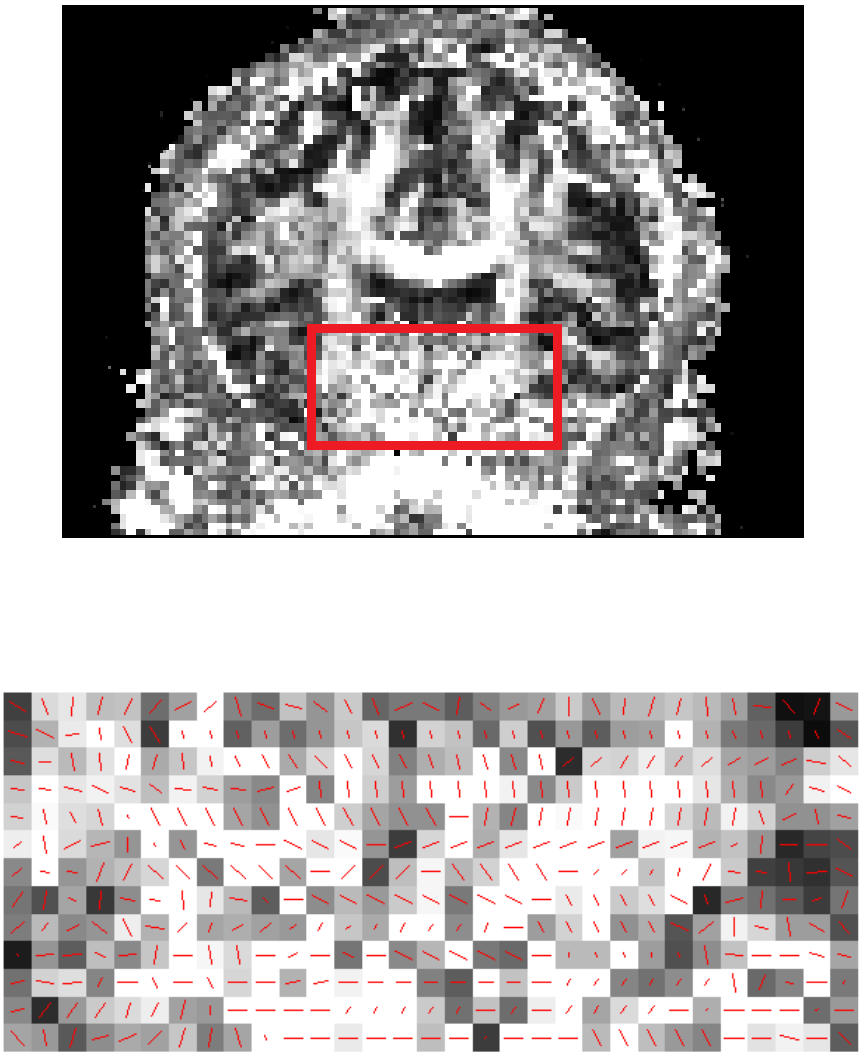


Figure 3.9 A slice of DTI images of a human brain in coronal view and its enlarge section.

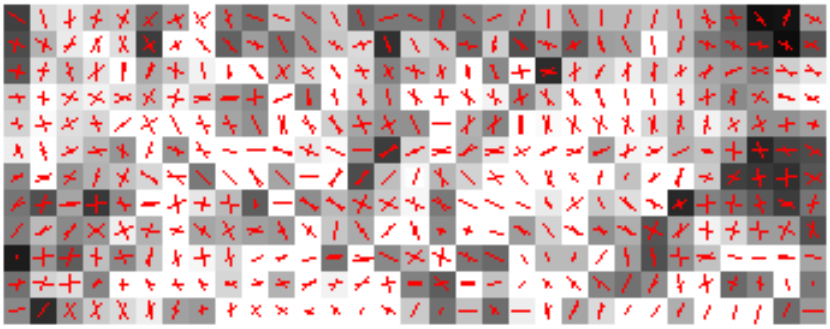
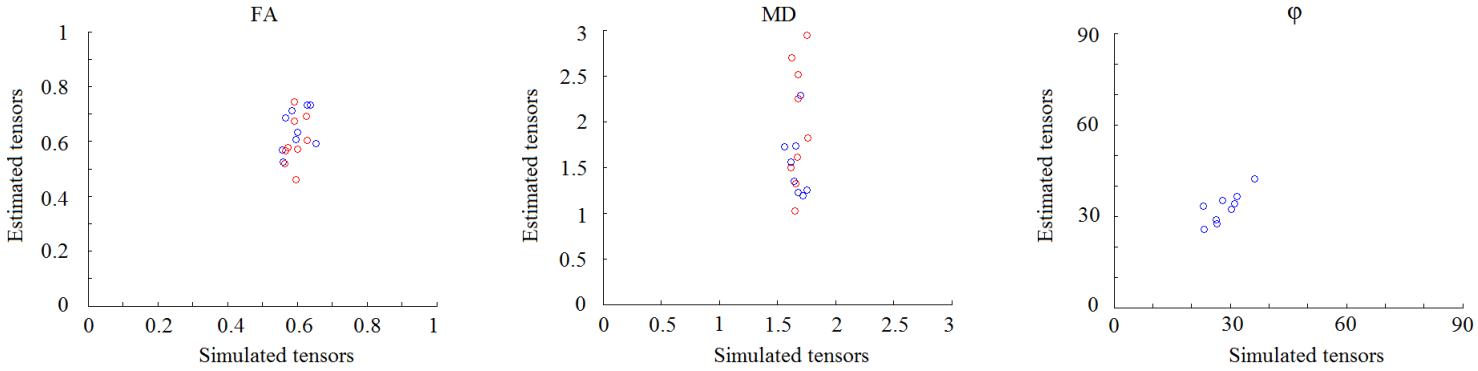


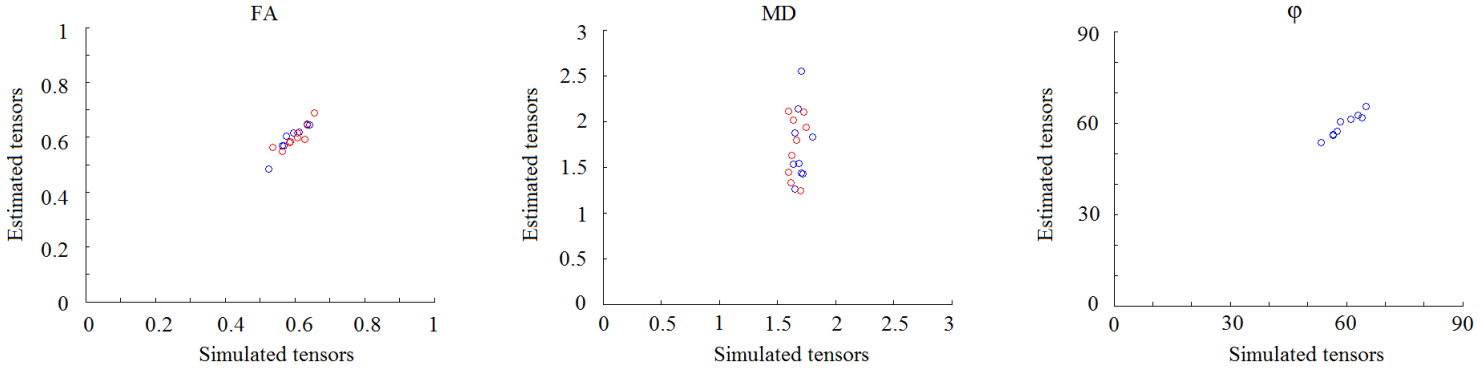
Figure 3.10 The FA and v_1 map of estimated tensors using BN algorithm. The model selection is performed at each voxel.

diffusion tensors with model selection available. The starting point and the phase portrait have been discussed and the procedure of the algorithm is stated in the framework. In the simulation studies, we simulated tensors fields that consist of single and double tensors at the crossing regions. We have examined the dependence of the properties of tensors and also the goodness of fit. The FA, MD and orientations of the estimated double tensors are shown in figures and tables. A real DTI image from a healthy human brain in coronal view is presented. The region of pons and corpus callosum is enlarged and the Bayesian models and BN algorithm are performed to estimate the diffusion tensors. Finally, extra Gaussian noise is added to such DTI image. We estimate the double tensors again and compare the results between raw data and noisy data.

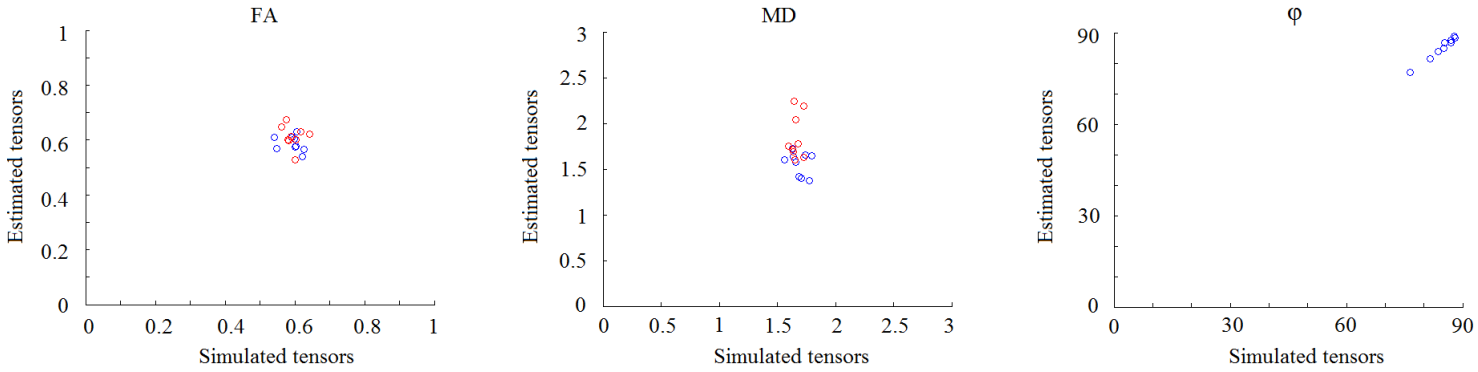
The results show that the Bayesian single and double model with neighbouring information as priors can identify the double tensors effectively and are robust. The Bayesian model is more accurate when two fibre orientations are distinct. When the two fibres orientation is similar, i.e. the angle of the principal direction is small, there may be a large variation of FA, MD of the estimated tensors. Especially, the determinant and MD tend to be larger than its true value. The results are optimal when the FA values are large, i.e. the tensors are more isotropic. When the two tensors have similar MD, the Bayesian models also perform better.



(a)Tensors $\mathbf{D}_1$ (blue) and $\mathbf{D}_2$ (red) are crossing at 30°

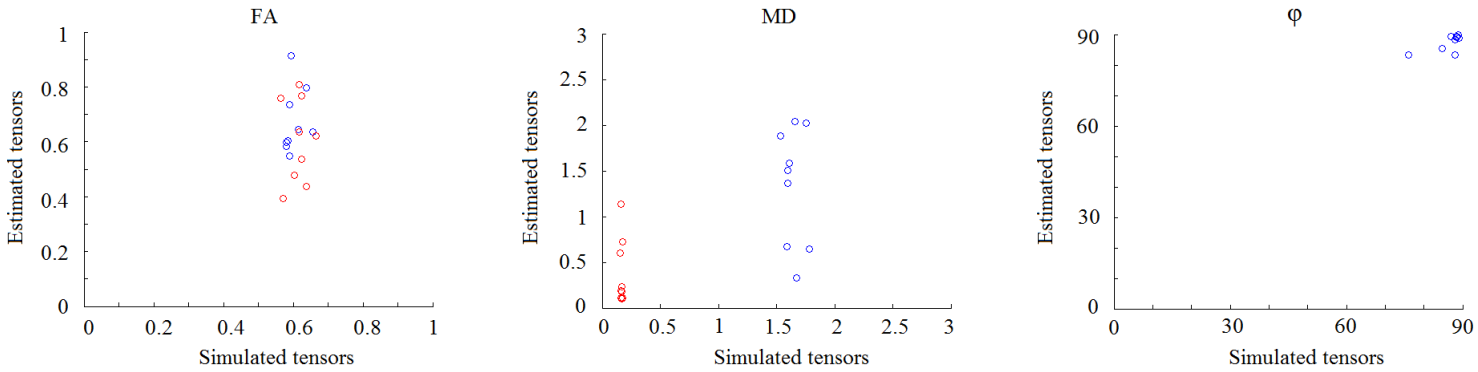


(b)Tensors $\mathbf{D}_1$ (blue) and $\mathbf{D}_2$ (red) are crossing at 60°

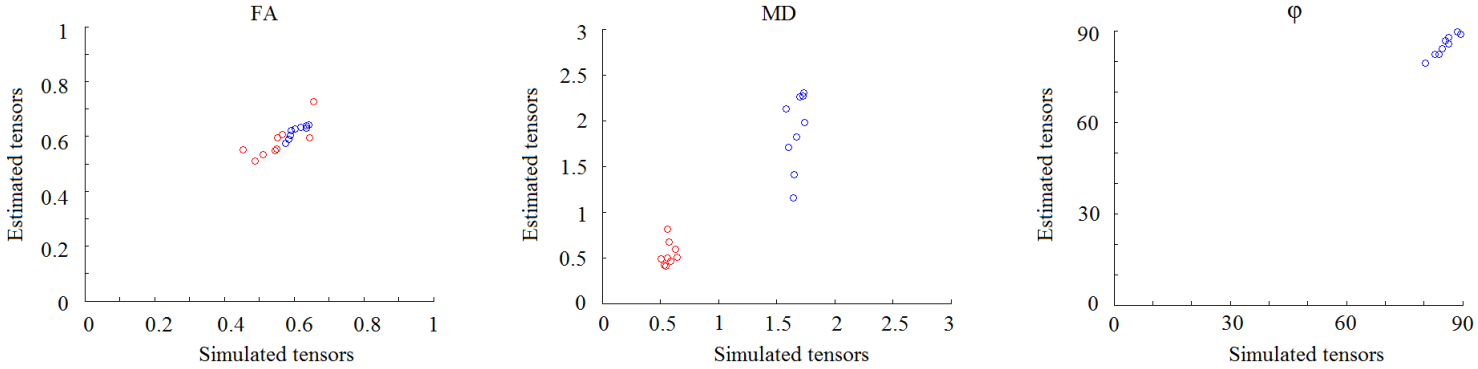


(c)Tensors $\mathbf{D}_1$ (blue) and $\mathbf{D}_2$ (red) are crossing at 90°

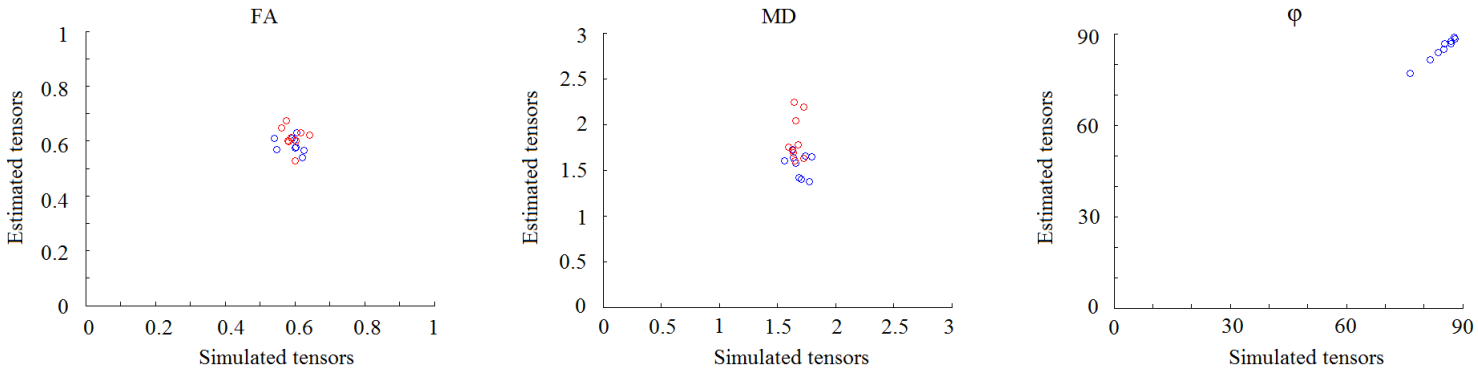
Figure 3.11 The scatter plots of FA, MD and orientations between simulated tensors and estimated tensors between two diffusion tensors crossing at 30°, 60° and 90°.



(a) $MD(\mathbf{D}_1) = 10MD(\mathbf{D}_2)$; ($\mathbf{D}_1$: blue, $\mathbf{D}_2$: red)

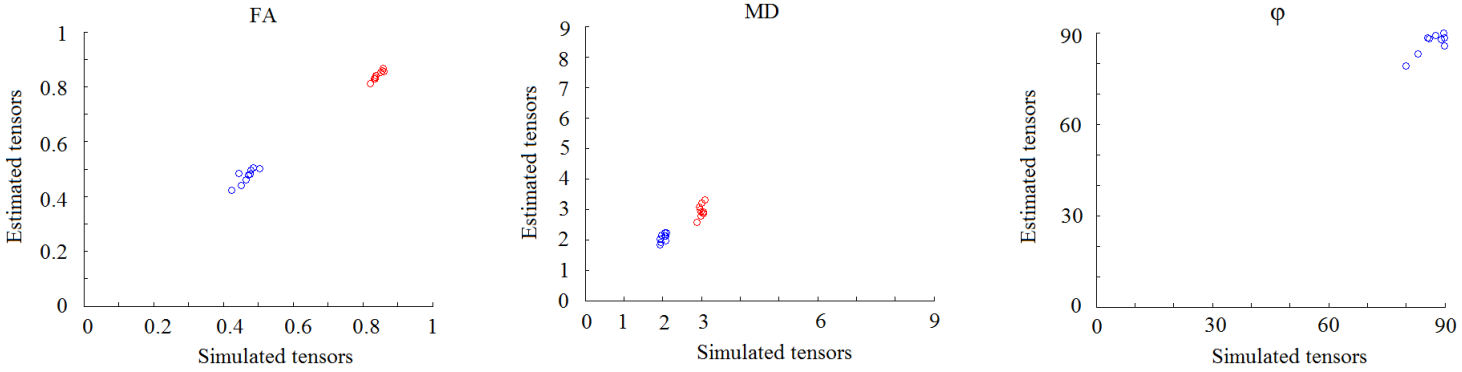


(b) $MD(\mathbf{D}_1) = 3MD(\mathbf{D}_2)$; ($\mathbf{D}_1$: blue, $\mathbf{D}_2$: red)

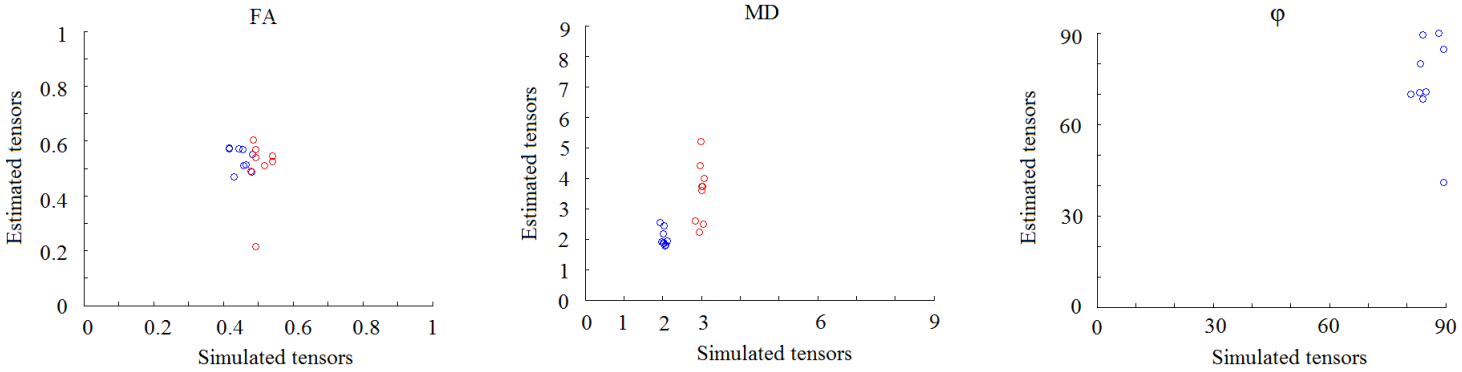


(c) $MD(\mathbf{D}_1) = MD(\mathbf{D}_2)$; ($\mathbf{D}_1$: blue, $\mathbf{D}_2$: red)

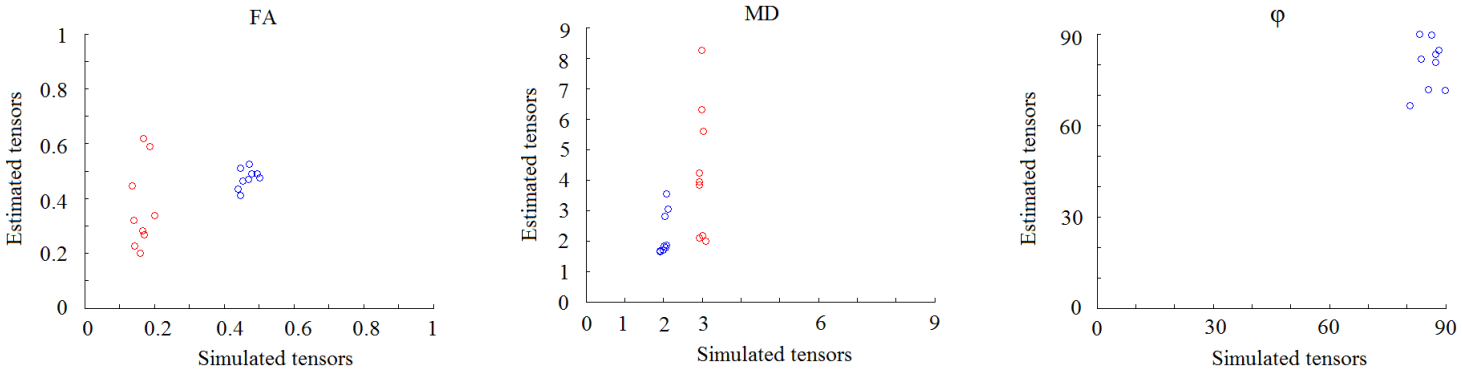
Figure 3.12 The scatter plots of FA, MD and angle between simulated tensors and estimated tensors if a tensor has the MD size of 10, 3 times and the same size of another tensor.



(a) $\mathbf{D}_2$ is a linear tensor; $\mathbf{D}_1$: blue, $\mathbf{D}_2$: red



(b) $\mathbf{D}_2$ is a planar tensor; $\mathbf{D}_1$: blue, $\mathbf{D}_2$: red



(c) $\mathbf{D}_2$ is a spherical tensor; $\mathbf{D}_1$: blue, $\mathbf{D}_2$: red

Figure 3.13 The scatter plots of FA, MD and angle between simulated tensors and estimated tensors if one tensor is a linear, planar or spherical tensor.

Chapter 4

Non-Euclidean Median Tensors Estimation

4.1 Introduction

In Chapter 2, we considered robust tensor estimation of a single tensor and multiple fibres problems. Some constraints were added ensuring the symmetric and positive semi-definite properties of tensor in case of the presence of noise arising in the measurement of diffusion attenuations corrupting these properties. The collection of tensors is known as a *tensor space*, or a *tensor field*. Fletcher and Joshi noted that the tensor field is actually a Riemannian manifold, rather than a Euclidean space due to the positive semi-definite symmetric tensors [29].

Using Euclidean distance in DTI may result in non-positive semi-definite matrices [25]. Thus it is natural to use non-Euclidean measurement techniques to estimate diffusion tensors. Due to the fact of possible outliers occurring in the experiments we mentioned in Section 2.2, e.g. thermal noise, spatially and temporally varying artifacts, such as subject motion, cardia pulsation and system instabilities, the obvious representative, the mean covariance tensor, may be easily affected by outliers [90, 30], i.e. it lacks robustness. In order to reduce the effect of possible outliers to a desirable level, following the idea of the previous chapter, median covariance matrices are considered in this chapter. Furthermore, the non-Euclidean nature of the tensor metric field is also considered. The geometric median has long been

used in Euclidean space, its use with general Riemannian manifold was first proposed by Fletcher [30].

In this chapter, the main focus is to define median diffusion tensors using non-Euclidean metrics. Our methods are applied to smooth and interpolate tensor fields. Comparisons are made between Euclidean mean, power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes techniques. Simulation studies and diffusion tensor imaging of a healthy human brain are considered.

4.1.1 Tensors operations

It is important to define the general tensor operations first, including power, square root, logarithm and exponential of a tensor. For any positive semi-definite tensor **D**, the usual spectral decomposition is as follows

$$D = U\Lambda U^T = U \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} U^T \tag{4.1}$$

with $U \in O(3)$ an orthogonal matrix and Λ diagonal with strictly positive entries.

The matrix power of tensor **D** to the power a, where $a \neq 0$, is

$$D^a = U \begin{pmatrix} \lambda_1^a & 0 & 0 \\ 0 & \lambda_2^a & 0 \\ 0 & 0 & \lambda_3^a \end{pmatrix} U^T \tag{4.2}$$

For $a = \dfrac{1}{2}$, we have the matrix square root:

$$D^{\frac{1}{2}} = U \begin{pmatrix} \lambda_1^{\frac{1}{2}} & 0 & 0 \\ 0 & \lambda_2^{\frac{1}{2}} & 0 \\ 0 & 0 & \lambda_3^{\frac{1}{2}} \end{pmatrix} U^T \tag{4.3}$$

The matrix logarithm of **D** is:

$$\log D = U \begin{pmatrix} \log \lambda_1 & 0 & 0 \\ 0 & \log \lambda_2 & 0 \\ 0 & 0 & \log \lambda_3 \end{pmatrix} U^T \tag{4.4}$$

and the exponential of a tensor **D** is defined as:

$$\exp D = U \begin{pmatrix} \exp \lambda_1 & 0 & 0 \\ 0 & \exp \lambda_2 & 0 \\ 0 & 0 & \exp \lambda_3 \end{pmatrix} U^T \tag{4.5}$$

These matrix operations ensure the results are also positive semi-definite. A tensor **D** has a unique symmetric matrix logarithm and conversely, each symmetric matrix is associated to a tensor by the exponential.

4.1.2 Euclidean and non-Euclidean metrics

We review the Euclidean distance and non-Euclidean distances. For any positive semi-definite tensor D_1 and D_2, the Euclidean distance between two tensors is the Euclidean norm of the difference of the two tensors, written as

$$d_E(D_1, D_2) = \|D_1 - D_2\| \tag{4.6}$$

An important non-Euclidean metric is the power Euclidean metric. The power Euclidean metric is defined for any pair of positive semi-definite diffusion tensors as

$$d_A(D_1, D_2) = \frac{1}{a}\|D_1^a - D_2^a\| \tag{4.7}$$

The Euclidean metric is a special form of power Euclidean metric where $a = 1$. Another special case is $a = \frac{1}{2}$, we have the root Euclidean metric. It defined as

$$d_H(D_1, D_2) = \|D_1^{\frac{1}{2}} - D_2^{\frac{1}{2}}\| \tag{4.8}$$

Note $d_H = \frac{1}{2} d_A(a = \frac{1}{2})$.

Another important non-Euclidean metric is the log-Euclidean distance. It is the Euclidean distance between the logarithm of diffusion tensors and defined as

$$d_L(D_1, D_2) = \| \log D_1 - \log D_2 \| \tag{4.9}$$

As the tensor field is a Riemannian symmetric manifold, it is natural to use another logarithm-based metric, the Riemannian metric

$$d_R(D_1, D_2) = \| \log D_1^{-\frac{1}{2}} D_2 D_1^{-\frac{1}{2}} \| \tag{4.10}$$

The Procrustes analysis is another approach in DTI study. Full ordinary Procrustes analysis is a form of statistical shape analysis used to match two objects with similarity transformations in terms of translation, rotation and uniform scale [26]. For two diffusion tensors D_1 and D_2, we are interested with scale transformation when match D_2 to D_1. We use a reparameterisation to ensure the positive semi-definiteness of diffusion tensor in the Procreates analysis, where Q_1 and Q_2 are 3×3 real matrices.

$$\begin{aligned} D_1 &= Q_1 Q_1^T \\ D_2 &= Q_2 Q_2^T \end{aligned} \tag{4.11}$$

The singular value decomposition of $Q_1^T Q_2$ is the product of two unitary matrices V and W^T and a matrix Δ with non-negative real number on the diagonal, the diagonal entries of Δ are known as the *singular values* of $Q_1^T Q_2$,

$$Q_1^T Q_2 = V \Delta W^T \tag{4.12}$$

Therefore a 3×3 rotation matrix $R \in O(3)$ is

$$R = W V^T \tag{4.13}$$

and the Procrustes size-and-shape distance is defined as

$$d_P(D_1, D_2) = \|Q_1 - Q_2 R\| \tag{4.14}$$

4.1.3 Fréchet Mean and mean covariance tensors

Since the space of diffusion tensors is non-Euclidean, it is more natural to estimate the mean covariance diffusion tensors using non-Euclidean metrics. The Fréchet mean covariance matrix defines a mean diffusion tensor T in a non-Euclidean space [31]

$$\hat{T} = \arg\inf_T \frac{1}{2} \int d(D,T)^2 f(D) dD \tag{4.15}$$

where $f(D)$ is the probability distribution for a 3×3 matrix $\mathbf{D}$ and $d(D,T)$ is a non-Euclidean metric of $\mathbf{D}$ and T.

We consider a sample of N semi-positive definite diffusion tensors $D_1, \cdots, D_N$. Assume that D_n are independently and identically distributed and $d(,)$ is the distance function. Thus the sample Fréchet Mean can be calculated by

$$\hat{T} = \arg\inf_T \sum_{n=1}^{N} d(D_i, T)^2 \tag{4.16}$$

While processing the tensor field in DTI, Euclidean mean tensor is often considered:

$$\hat{T}_E = \arg\inf_T d_E(D_i, T)^2 = \frac{1}{N} \sum_{i=1}^{N} D_i \tag{4.17}$$

The mean covariance tensors in terms of different non-Euclidean metrics are expressed in the Table 4.1 [25].

Name	Metric	Mean tensor
Power Euclidean	$\frac{1}{a}\|\|D_1^a - D_2^a\|\|$	$(\frac{1}{N}\sum_{n=1}^{N} D_n^a)^{\frac{1}{a}}$
Root Euclidean	$\|D_1^{\frac{1}{2}} - D_2^{\frac{1}{2}}\|$	$(\frac{1}{N}\sum_{n=1}^{N} D_n^{\frac{1}{2}})(\frac{1}{N}\sum_{n=1}^{N} D_n^{\frac{1}{2}})^T$
Log-Euclidean	$\|\log D_1 - \log D_2\|$	$\exp\{\frac{1}{N}\sum_{n=1}^{N} \log D_n\}$
Riemannian	$\|\log D_1^{-\frac{1}{2}} D_2 D_1^{-\frac{1}{2}}\|$	$\arg\inf \sum_{n=1}^{N} \|\log (D_1^{-\frac{1}{2}} T D_1^{-\frac{1}{2}})\|^2$
Procrustes	$\inf_{R\in O(3)} \|Q_1 - Q_2 R\|$	$\arg\inf \sum_{n=1}^{N} \inf \|\|Q_1 - Q_2 R\|\|$

Table 4.1 Summaries of metrics and mean covariance tensor estimators.

4.2 Non-Euclidean estimators of median covariance tensors

In two-dimensional space, the median is the point that separates the data set into equal halves on both sides (if n is odd) and is any point on the line segment connecting the two middle points (if n is even). In a three-dimensional space or metric space, L_1 median is defined to be any point which minimises the sum of Euclidean distances to all points in the data set. In order to reduce the effect of the existence of outliers, we use the geometry median in covariance tensor calculation. Here we define the median as the point that generates the least sum of chosen distances to every other matrix in the tensor field.

$$\text{Median} = \arg\inf_{T} \sum_{n=1}^{N} d_*(D_n, T) \tag{4.18}$$

where w_n are the weights, D_n are the tensors, T is the required median tensor and d_* is the non-Euclidean distance function.

Thus the median covariance tensor is the matrix given the least $F(T)$, where

$$F(T) = \sum_{n=1}^{N} d(D_n, T) \tag{4.19}$$

In other words, T is required when the function $F(T)$ is minimised. Thus, we solve the gradient function $\nabla F(T) = \mathbf{0}$.

The Euclidean median can be estimated by Weiszfeld's algorithm [18, 56], which uses iteratively re-weighted least squares. This algorithm converges in most cases. When the estimates fail on one of the given points, the algorithm requires modification to achieve convergence [107]. Recent research by Cohen et al. [21] has shown the computation of geometric median in nearly linear time. In this section, we will define the median covariance tensor in terms of different non-Euclidean distances and their corresponding gradient functions.

4.2.1 Estimators with reparameterisation

An approach of analysing positive definite tensors is the reparameterisation of diffusion tensors. The median tensor with respect to power Euclidean distance is given as

$$T_A = \arg\inf_T \sum_{n=1}^{N} \frac{1}{a} ||D_n^a - T^a|| \tag{4.20}$$

Thus the corresponding gradient function ∇T_A is

$$\begin{aligned}
\nabla T_A &= \frac{\partial \sum_{n=1}^{N} \frac{1}{a} ||D_n^a - T^a||}{\partial T} \\
&= \sum_{n=1}^{N} \frac{\partial ||D_n^a - T^a||}{a \partial T} \\
&= \sum_{n=1}^{N} \frac{2(D_n^a - T^a)(-aT^{a-1})}{2\,||D_n^a - T^a||} \\
&= \sum_{n=1}^{N} \frac{aT^{a-1}(T^a - D_n^a)}{||D_n^a - T^a||}
\end{aligned}$$

A usual setting is $a = 2$, we have the first derivative

$$\nabla T_A = \sum_{n=1}^{N} \frac{2T(T^2 - D_n^2)}{(T^2 - D_n^2)} \tag{4.21}$$

A special form of power Euclidean median tensor is the root Euclidean median tensor ($a = \frac{1}{2}$):

$$T_H = \arg\inf_T \sum_{n=1}^{N} \|D_n^{\frac{1}{2}} - T^{\frac{1}{2}}\| \tag{4.22}$$

and the gradient function of median tensor in terms of root Euclidean metric is as follows:

$$\nabla\, T_H = \sum_{n=1}^{N} \frac{(\sqrt{T} - \sqrt{D_n})}{2\sqrt{T}\|D_n^{\frac{1}{2}} - T^{\frac{1}{2}}\|} \tag{4.23}$$

4.2.2 Logarithm-based median estimators

Considering the log-Euclidean distance, we obtain the median tensor of the log-Euclidean metric as follows

$$T_L = \arg\inf_T \sum_{n=1}^{N} ||\log D_n - \log T|| \tag{4.24}$$

The gradient function of log-Euclidean median tensor is

$$\begin{aligned}
\nabla T_P &= \frac{\partial \sum_{n=1}^{N} \|\log D_n - \log T\|}{\partial T} \\
&= \sum_{n=1}^{N} \frac{2(\log D_n - \log T)}{2\|\log D_n - \log T\|}(-\frac{1}{T}) \\
&= \sum_{n=1}^{N} \frac{\log D_n - \log T}{T\|\log D_n - \log T\|}
\end{aligned}$$

An alternative logarithm-based metric is the Riemannian metric. The median covariance tensor of Riemannian metric is

$$T_R = \arg\inf_T \sum_{n=1}^{N} \| \log D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}} \| \tag{4.25}$$

and the corresponding gradient function is

$$\begin{aligned} \nabla T_R &= \frac{\partial \sum_{n=1}^{N} \| \log D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}} \|}{\partial T} \\ &= \sum_{n=1}^{N} \frac{2 \log D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}}}{2 \| \log D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}} \|} \times \frac{1}{D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}}} \\ &= \sum_{n=1}^{N} \frac{D_n \log D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}}}{T \| \log D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}} \|} \end{aligned}$$

4.2.3 Procrustes estimator

The Procrustes analysis matches two objectives in terms of translation, rotation and scale. Here we use a discrete method to calculate the Procrustes median tensor estimator. The median tensor of Procrustes size-and-shape is

$$\hat{T}_P = \hat{Q}_P \hat{Q}_P{}^T \tag{4.26}$$

where

$$\hat{Q}_P = \frac{1}{N} \sum_{n=1}^{N} Q_n R_n \tag{4.27}$$

where $D_n = Q_n Q_n^T$.

4.3 Weights

In the previous sections we assume the weights w_i are equal,

$$w_i = \frac{1}{N} \tag{4.28}$$

In this section, we explore weighted non-Euclidean median tensors.

4.3.1 Weighting functions

The non-negative weights in diffusion tensor calculation need to satisfy:

$$\sum w_n = 1, \qquad n = 1, \cdots N. \tag{4.29}$$

A simple set of weights is given by

$$w_n = \frac{d_n^{-1}}{\sum_{j=1}^{N} d_j^{-1}}, \qquad m = 1, \cdots, N. \tag{4.30}$$

where d_n is the Euclidean distance from the voxel of weighted median to the nth voxel with D_n.

An alternative system of weights are determined using the exponential weight function

$$w_n = \frac{\exp(-Ad_n^2) + B}{\sum_{m=1}^{N} \exp(-Ad_m^2) + B}, \qquad m = 1, \cdots, N. \tag{4.31}$$

where A and B are arbitrary non-negative constants.

The exponential weights are greater for smaller values of d_n.

4.3.2 Weighted metrics and median covariance tensors

We have discussed several choices of non-weighted metrics between diffusion tensors in DTI studies. The weighted non-Euclidean metrics and median tensors estimators T_A, T_H, T_L, T_R and T_P are listed in the Table 4.2. For completeness, the Euclidean metric and Euclidean

mean tensor T_E are also listed in Table 4.2. The Euclidean mean tensor estimators T_E and Procrustes size-and-shape median tensor estimator T_P are calculated using discrete arithmetic operations. The other four non-Euclidean median tensor estimators are computed using the gradient descent algorithm, which is described in the Section 4.4.

Name	Estimator	Weighted metrics	Weighted tensor estimators
Euclidean	T_E	$w_n\|D_1 - D_2\|$	$\arg\inf_T \sum_{n=1}^{N} d_*(D_n, T)$
Power Euclidean	T_A	$\frac{w_n}{a}\|D_1^a - D_2^a\|$	$\arg\inf \sum_{n=1}^{N} \frac{w_n}{a}\|D_n^a - T^a\|$
Root Euclidean	T_H	$w_n\|D_1^{\frac{1}{2}} - D_2^{\frac{1}{2}}\|$	$\arg\inf \sum_{n=1}^{N} w_n \|D_n^{\frac{1}{2}} - T^{\frac{1}{2}}\|$
Log-Euclidean	T_L	$w_n\|\log D_1 - \log D_2\|$	$\arg\inf \sum_{n=1}^{N} w_n\|\log D_n - \log T\|$
Riemannian	T_R	$w_n\|\log D_1^{-\frac{1}{2}} D_2 D_1^{-\frac{1}{2}}\|$	$\arg\inf \sum_{n=1}^{N} w_n\|\log D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}}\|$
Procrustes	T_P	$\inf_R w_n\|Q_1 - Q_2 R\|$	$\hat{Q}_s \hat{Q}_s^T$

Table 4.2 Summaries of Euclidean and non-Euclidean estimators, weighted metrics and weighted tensor estimators.

4.4 Algorithm

We use the gradient-descent algorithm, which is a first-order optimisation algorithm, to obtain the median estimators. It ensures convergence of the function [Pennec et al.]. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative gradient (or of the approximate gradient) of the function at the current point. If instead one takes steps proportional to the positive gradient, one approaches a local maximum of that function; the procedure is then known as gradient ascent.

The objective function is $\nabla f(\mathbf{T}) = \mathbf{0}$, where

$$f(\mathbf{T}) = \sum_{n=1}^{N} w_n d(D_n, T) \tag{4.32}$$

The median covariance tensor is the matrix given the least $F(T)$, where

$$F(T) = \sum_{n=1}^{N} d(D_n, T) \tag{4.33}$$

In other words, T is acquired when the function $F(T)$ is minimised. Thus, we solve the gradient function $\nabla F(T) = \mathbf{0}$. The gradient functions for the weighted median estimation using the power Euclidean, root Euclidean, log-Euclidean and Riemannian metrics are illustrated respectively as follow:

$$f_A = \nabla T_A = \sum_{n=1}^{N} \frac{a w_n T^{a-1}}{||D_n^a - T^a||}(T^a - D_n^a) \tag{4.34}$$

$$f_H = \nabla T_H = \sum_{n=1}^{N} \frac{w_n(\sqrt{T} - \sqrt{D_n})}{2\sqrt{T}||D_n^{\frac{1}{2}} - T^{\frac{1}{2}}||} \tag{4.35}$$

$$f_L = \nabla T_L = \sum_{n=1}^{N} \frac{w_n(\log T - \log D_n)}{T\|\log D_n - \log T\|} \tag{4.36}$$

$$f_R = \nabla T_R = \sum_{n=1}^{N} \frac{w_n D_n \log(D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}})}{T\|\log(D_n^{-\frac{1}{2}} T D_n^{-\frac{1}{2}})\|} \tag{4.37}$$

We use the Euclidean mean covariance tensor T_E as an initial value of the non-Euclidean estimators. The algorithm is taking a step $\alpha > 0$ in the reduction of $\nabla F(T)$ at each iteration, where α is usually very small. Taking the log-Euclidean estimator as an example, the algorithm is as follows:

1. **Initial setting:** $f_L \leftarrow \infty$
2. $\hat{T}_L$ from precious iteration: $T_i \leftarrow T_E$
3. **while** $|f_L| >$ tolerance **do**
4. **for** $i = 1$ to N **do**
5. $$f_L = \sum_{n=1}^{N} \frac{(\log T_i - \log D_n)}{T_i d_L(D_n, T_i)}$$
6. **end for**

7. $\hat{T}_L$ from current iteration: $T_{i+1} = T_i - \alpha f_L$
8. $T_i \leftarrow T_{i+1}$
9. **end while**
10. $\hat{T}_L \leftarrow T_{i+1}$
11. **return** $\hat{T}_L$

4.5 Simulation studies

We carry out a simulation study to compare the median covariance tensor estimators in terms of different non-Euclidean methods. We consider the problem of estimating a population covariance matrix D_0 from a random sample 5×5 covariance matrices. The tensor D_0 is:

$$D_0 = \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.3 \end{pmatrix} \tag{4.38}$$

The tensor D_0 has determinant $det(D_0) = 0.162$, mean diffusivity $MD = 0.6$ and fractional anisotropy $FA = 0.4629$.

The random sample is generated as follows. Let D_n be a random matrix with independent and identically normally distributed entries with $E[(D_n)_{jl}] = 0$, $var[(D_n)_{jl}] = \sigma^2$, $n = 1, \cdots, N$, $j = 1, \cdots, k, l = 1, \cdots, k$. We consider the Gaussian square root error model with three different variances: 0.05, 0.1 and 0.2.

The median covariance tensors are calculated in terms of log-Euclidean distance(T_L), power Euclidean distance(T_P) and root Euclidean distance(T_H). We estimate the median tensors of 3×3, 4×4 and 5×5 tensors respectively. When $n = 3$, we use the tensors in the top three rows and left three columns. When $n = 4$, we use the tensors in the top four rows and left four columns. When $n = 5$, all the tensors in the graph are used in the calculation. The results are compared in the Tables 4.3, 4.4 and 4.5. These tables compare the determinant, mean diffusivity and fractional anisotropy value of the three estimated median tensors.

		det(T)	MD	FA
$n=3$	T_E	0.7686	2.0786	0.4517
	T_A	0.2625	0.6955	0.4463
	T_L	0.2546	0.6911	0.4533
	T_H	0.2551	0.6916	0.4537
	T_R	0.2548	0.6918	**0.4545**
	T_P	**0.2525**	**0.6903**	0.4537
$n=4$	T_E	0.9628	2.7150	0.4504
	T_A	**0.2437**	0.6780	0.4411
	T_L	0.2390	0.6757	0.4471
	T_H	0.2393	0.6758	0.4470
	T_R	0.2394	**0.6753**	0.4447
	T_P	0.2372	0.6762	**0.4527**
$n=5$	T_E	1.2025	3.3924	0.4550
	T_A	0.2482	0.6817	0.4453
	T_L	0.2390	0.6760	0.4528
	T_H	0.2401	0.6769	0.4526
	T_R	0.2392	**0.6756**	0.4508
	T_P	**0.2369**	0.6759	**0.4574**

Table 4.3 Measures of tensors properties with $\sigma = 0.05$.

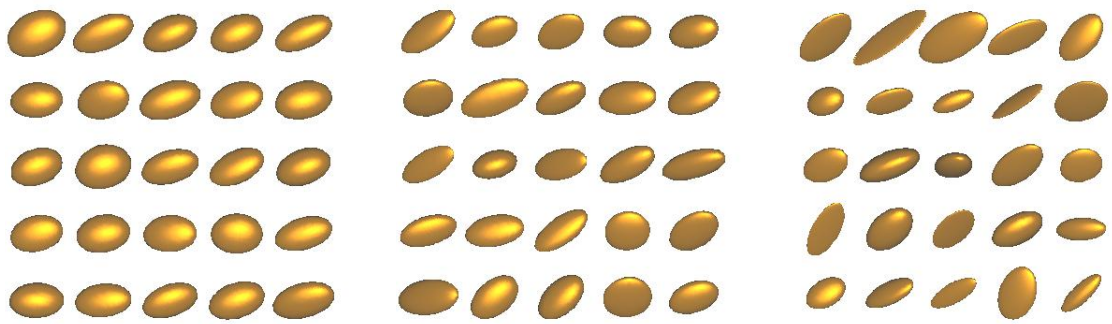


Figure 4.1 Simulated metric space with variance 0.05(left), 0.1(middle) and 0.2(right).

In Tables 4.3, 4.4 and 4.5, the best results, i.e. those are closest to D_0, are highlighted in bold. It is obvious that the results using non-Euclidean distances are better than the results calculated from the Euclidean distance, not only the properties of tensors but also the distances between the original tensor D_0.

		det(T)	MD	FA
$n=3$	T_E	0.7541	2.1647	0.5437
	T_A	0.2664	0.7265	**0.5212**
	T_L	**0.2158**	**0.6977**	0.5586
	T_H	0.2269	0.7046	0.5510
	T_R	**0.2158**	0.6983	0.5600
	T_P	0.2289	0.7093	0.5566
$n=4$	T_E	1.0409	2.9103	0.5330
	T_A	0.2883	0.7405	**0.5088**
	T_L	**0.2351**	**0.708**2	0.5414
	T_H	0.2441	0.7150	0.5374
	T_R	**0.2351**	0.7095	0.5440
	T_P	0.2422	0.7167	0.5439
$n=5$	T_E	1.4192	3.7013	0.5155
	T_A	0.3074	0.7473	**0.4877**
	T_L	**0.2633**	**0.7184**	0.5084
	T_H	0.2724	0.7255	0.5053
	T_R	0.2635	0.7188	0.5091
	T_P	0.2643	0.7291	0.5273

Table 4.4 Measures of tensors properties with $\sigma = 0.1$.

All non-Euclidean distance estimators obtain reasonably good results. When the variance and scale are small, the discrete Procrustes method obtains the best results. As the scale gets larger, the log-Euclidean estimators have the closest value to the D_0 when the variance of tensors is 0.05. When the variance gets larger, 0.1 and 0.2, the Riemannian estimators and the log-Euclidean estimators both perform better than the other estimators. The determinant of power Euclidean tensor is much less that of than the other two tensors. The root Euclidean tensor has the least mean diffusivity whilst the power Euclidean tensor has the greatest value. The FA values of the median tensor are similar between three methods.

In the respect of FA value, which is in the range $[0,1]$, there is not much difference between the different estimators in Table 4.3. The Procrustes estimators perform slightly better. In Tables 4.4 and 4.5, where there is a larger variance, the power Euclidean estimators have an advantage in estimating the FA value over others, and they have the closest value to D_0.

Additionally, the non-Euclidean distances between each two estimated tensors are shown in the Tables 4.6, 4.7and 4.8.The estimated risk using Stein loss [46], which is given by

$$L(D_1, D_2) = \text{trace}(D_1 D_2^{-1}) - \log(\det(D_1 D_2^{-1})) - k, \tag{4.39}$$

In Tables 4.6, 4.7 and 4.8, we list all the distances between all the estimators and the original tensor D_0. All the non-Euclidean estimators produce smaller estimators of distance than the Euclidean estimators. Whilst there is little variation amongst the estimators using non-Euclidean methods, there is considerable variation of those using Euclidean methods. Generally speaking, the log-Euclidean and root Euclidean median tensor estimators have better results than other estimators. The power Euclidean estimators provide the largest distances among the non-Euclidean methods. It is interesting to point out that the root Euclidean estimators always produce the smaller estimators that in root Euclidean distance and Riemannian distances.

		det(T)	MD	FA
$n=3$	T_E	1.8672	2.9187	0.5539
	T_A	0.6416	0.9509	**0.4880**
	T_L	0.4163	**0.8638**	0.5633
	T_H	0.4713	0.8934	0.5553
	T_R	**0.4162**	0.8712	0.5775
	T_P	0.5167	0.8316	0.5767
$n=4$	T_E	2.3274	3.8719	0.5708
	T_A	0.7284	1.0085	**0.5135**
	T_L	0.4382	**0.9014**	0.5886
	T_H	0.5025	0.9330	0.5750
	T_R	**0.4377**	0.9046	0.5929
	T_P	0.4833	0.9300	0.5963
$n=5$	T_E	2.6432	4.6762	0.5619
	T_A	0.4735	0.8915	**0.5225**
	T_L	0.3415	**0.8273**	0.5622
	T_H	0.3735	0.8466	0.5545
	T_R	**0.3409**	0.8309	0.5701
	T_P	0.4441	0.9011	0.5868

Table 4.5 Measures of tensors properties with $\sigma = 0.2$.

size	Estimators	d_E	d_A	d_L	d_H	d_R	d_P	Stein
$n=3$	T_E	2.7824	4.9604	1.8861	1.1177	1.8861	1.1657	2.9041
	T_A	0.2617	0.2058	0.4100	0.1581	0.4140	0.1154	0.0978
	T_L	0.2539	**0.2026**	0.3867	0.1586	0.4131	0.1223	0.0885
	T_H	0.2556	0.2031	0.3922	**0.1530**	**0.3962**	0.1100	0.0891
	T_R	**0.2517**	0.2034	**0.3772**	0.1247	0.4203	0.1616	**0.0850**
	T_P	0.2551	0.1997	0.3934	0.1535	0.3981	0.1523	0.0905
$n=4$	T_E	3.9547	8.6671	2.3191	1.4649	2.3191	1.5201	4.8310
	T_A	0.2248	0.1700	0.3608	0.1379	0.3655	0.1060	0.0749
	T_L	0.2222	0.1708	0.3473	0.1412	0.3721	0.1005	0.0708
	T_H	0.2219	0.1694	0.3498	**0.1350**	**0.3545**	0.1006	0.0703
	T_R	**0.2183**	**0.1681**	**0.3411**	0.1426	0.3776	0.1011	**0.0688**
	T_P	0.2286	0.1749	0.3548	0.1384	0.3604	0.1370	0.0731
$n=5$	T_E	5.2227	13.8945	2.6862	1.7959	2.6862	1.8556	7.0656
	T_A	0.2522	0.1850	0.4528	0.1623	0.4552	0.1616	0.1183
	T_L	0.2243	0.1758	0.3480	0.1397	0.3685	0.1384	**0.0706**
	T_H	0.2252	0.1764	0.3516	**0.1359**	**0.3557**	0.1348	0.0707
	T_R	**0.2205**	**0.1731**	**0.3429**	0.1408	0.3726	0.1394	0.0689
	T_P	0.2278	0.1781	0.3520	0.1371	0.3570	0.1358	0.0716

Table 4.6 Measures of efficiency with different distances ($\sigma = 0.05$).

size	Estimators	d_E	d_A	d_L	d_H	d_R	d_P	Stein
$n=3$	T_E	3.1411	6.2698	2.2433	1.2505	2.2485	1.2490	4.4772
	T_A	0.4825	0.3999	0.6985	0.2823	0.7117	0.2791	0.3186
	T_L	**0.4645**	**0.3853**	**0.6708**	0.2785	0.7108	0.2738	**0.2849**
	T_H	0.4689	0.3887	0.6766	**0.2740**	**0.6925**	0.2703	0.2908
	T_R	0.4675	0.3882	0.6783	0.2791	0.7117	0.2744	0.2886
	T_P	0.4826	0.4058	0.6817	0.2793	0.7003	0.2750	0.2988
$n=4$	T_E	4.5329	11.4008	2.7319	1.6427	2.7353	1.6416	7.2553
	T_A	0.4550	0.3877	0.6414	0.2624	0.6535	0.2594	0.2676
	T_L	**0.4278**	**0.3635**	**0.5961**	0.2531	0.6318	0.2492	**0.2271**
	T_H	0.4348	0.3694	0.6062	**0.2500**	**0.6210**	0.2466	0.2343
	T_R	0.4318	0.3686	0.6015	0.2549	0.6349	0.2510	0.2302
	T_P	0.4466	0.3823	0.6151	0.2554	0.6314	0.2516	0.2422
$n=5$	T_E	5.9994	18.1815	3.1601	2.0175	3.1631	2.0165	10.7581
	T_A	0.4317	0.3721	0.5955	0.2462	0.6049	0.2438	0.2255
	T_L	0.4034	0.3434	0.5491	0.2350	0.5816	0.2323	0.1936
	T_H	0.4095	0.3504	0.5596	**0.2332**	**0.5701**	0.2307	**0.1981**
	T_R	**0.4044**	**0.3450**	**0.5478**	0.2360	0.5840	0.2332	0.1933
	T_P	0.4565	0.3896	0.6308	0.2615	0.6465	0.2578	0.2586

Table 4.7 Measures of efficiency with different distances ($\sigma = 0.1$).

size	Estimators	d_E	d_A	d_L	d_H	d_R	d_P	Stein
$n=3$	T_E	4.6422	12.1718	2.7702	1.6663	2.7744	1.6649	7.8057
	T_A	0.8210	0.8362	1.0247	0.4406	1.0326	0.4384	0.7495
	T_L	**0.7387**	**0.7676**	**0.8643**	0.4178	0.9526	0.4142	**0.5323**
	T_H	0.7807	0.8150	0.9123	**0.4077**	**0.9250**	0.4046	0.5921
	T_R	0.7709	0.8177	0.8766	0.4270	0.9663	0.4234	0.5539
	T_P	0.8906	0.9860	0.9826	0.4514	0.9961	0.4481	0.7060
$n=4$	T_E	6.5643	22.1736	3.2353	2.1199	3.2393	2.1185	12.1359
	T_A	0.9796	1.0537	1.1473	0.5099	1.1556	0.5077	0.9820
	T_L	**0.8648**	**0.9237**	**0.9895**	**0.4604**	**1.0339**	0.4569	**0.7123**
	T_H	0.9069	0.9759	1.0311	0.4674	1.0439	0.4643	0.7821
	T_R	0.8769	0.9419	0.9947	0.4640	1.0398	0.4604	0.7238
	T_P	0.9370	1.0368	1.0305	0.4751	1.0446	0.4716	0.7897
$n=5$	T_E	8.0872	31.8369	3.5474	2.4489	3.5511	2.4476	15.6009
	T_A	0.7637	0.7402	0.9311	0.4104	0.9455	0.4069	0.6190
	T_L	**0.6877**	**0.6680**	**0.8528**	0.4260	0.9788	0.4227	**0.5048**
	T_H	0.7169	0.6869	0.8764	**0.3865**	**0.8918**	0.3829	0.5389
	T_R	0.7081	0.6940	0.8601	0.4274	0.9862	0.4241	0.5197
	T_P	0.8615	0.9225	0.9641	0.4419	0.9793	0.4383	0.6804

Table 4.8 Measures of efficiency with different distances ($\sigma = 0.2$).

4.5.1 Single tensor estimators with weights

The second simulation study is to compare the different weighted median covariance tensor estimators in terms of different non-Euclidean methods. We consider the problem of estimating a population covariance matrix D_0 from a random sample 5×5 covariance matrices. We focus on the tensor field in the middle of Figure 4.1. Here, we have the same tensor D_0:

$$D_0 = \begin{pmatrix} 0.9 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & 0.3 \end{pmatrix} \tag{4.40}$$

and the random samples are generated from the Gaussian square root error model with variance 0.1.

The tensor T_0 at the central voxel (with coordinates $[3,3]$) is chosen for each metric space. The weighted median covariance tensors are calculated in terms of Euclidean distance (T_E) and various non-Euclidean distances. We estimate the median tensors of 5×5 tensors, i.e. all the tensors in the graph are used in the calculation. The weights we use equal weight (Equation (4.28)), inverse Euclidean weight (Equation (4.30)) and exponential weight (Equation (4.31)). The determinant, MD, FA of the tensor T_0 at the central voxel and the estimators are compared in Table 4.9. The angles between the main directions ϕ of T_0 and estimators are also shown in the table. The best results are highlighted in terms of each weight setting.

The weighted estimators improve the accuracy of determinant, MD, FA and main orientations comparing to equal weight. The inverse Euclidean weight has an advantage of Euclidean mean, power and root Euclidean estimators in terms of determinant, MD and FA. In contrast, the exponential weight have preferable estimators of log-Euclidean, Riemannian, and Procrustes metrics. In terms of the orientation of main fibres, the exponential weight definitely performs the best.

		det(T)	MD	FA	ϕ
T_0		0.2106	0.7272	0.6194	33.01°
T_E	w_1	0.2838	0.7403	0.5155	28.38°
	w_2	**0.2820**	**0.7383**	**0.5177**	28.20°
	w_3	0.2823	0.7391	0.5176	**28.62**°
T_A	w_1	0.2926	0.7432	0.5056	28.45°
	w_2	**0.2903**	**0.7421**	**0.5109**	28.46°
	w_3	0.2925	0.7434	0.5081	**28.60**°
T_L	w_1	0.2663	0.7304	0.5271	28.28°
	w_2	0.2667	**0.7299**	0.5286	28.19°
	w_3	**0.2657**	**0.7299**	**0.5288**	**28.57**°
T_H	w_1	0.2838	0.7402	0.5156	28.38°
	w_2	**0.2820**	**0.7382**	**0.5177**	28.20°
	w_3	0.2822	0.7390	**0.5177**	**28.62**°
T_R	w_1	0.2647	0.7300	0.5288	28.24°
	w_2	0.2655	**0.7295**	0.5299	28.18°
	w_3	**0.2643**	**0.7295**	**0.5303**	**28.55**°
T_P	w_1	0.2643	0.7291	0.5273	28.46°
	w_2	0.2629	**0.7272**	0.5291	28.27°
	w_3	**0.2617**	0.7274	**0.5302**	**28.71**°

Table 4.9 Measures of tensors properties with $\sigma = 0.1$.

4.5.2 Smoothing

The median tensor estimators can be used to smooth the diffusion tensor space. We form a 10×10 tensor space with the four distinct tensors of four corners of Figure 4.2. A Gaussian noise of 0.5 is added on all the tensors as in the top left graph. We calculate the median tensor estimators using its neighbouring six voxels and itself, then use the estimators to replace each tensor. The results of FA and main diffusion direction v_1 maps are generated in Figure 4.3. The median images look much smoother than their original image and the directions of v_1 are more clearly illustrated the fibres in the brain. The log-Euclidean and Riemannian distances method are similar. They both remain abnormal boxes (with coordinates [10,7] and [6,10]), while the local variation for other methods ia smaller. The top right corner of the Riemannian method is slightly darker than others.

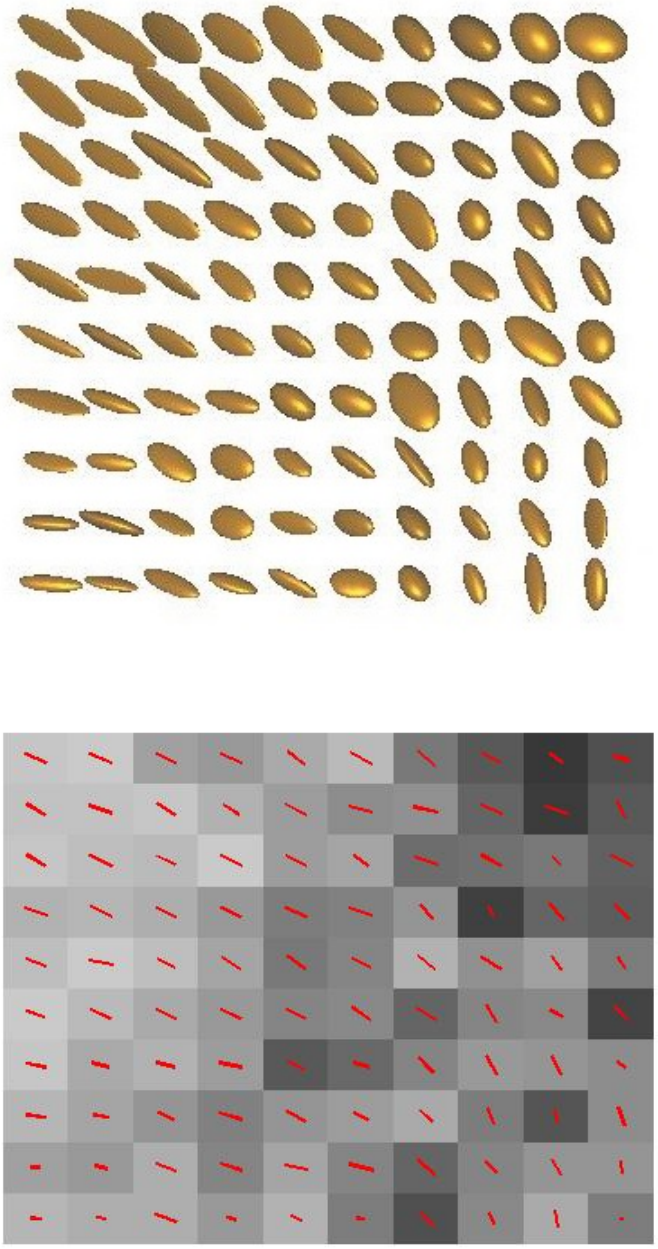


Figure 4.2 Simulated 10×10 tensors and its FA and v1 map.

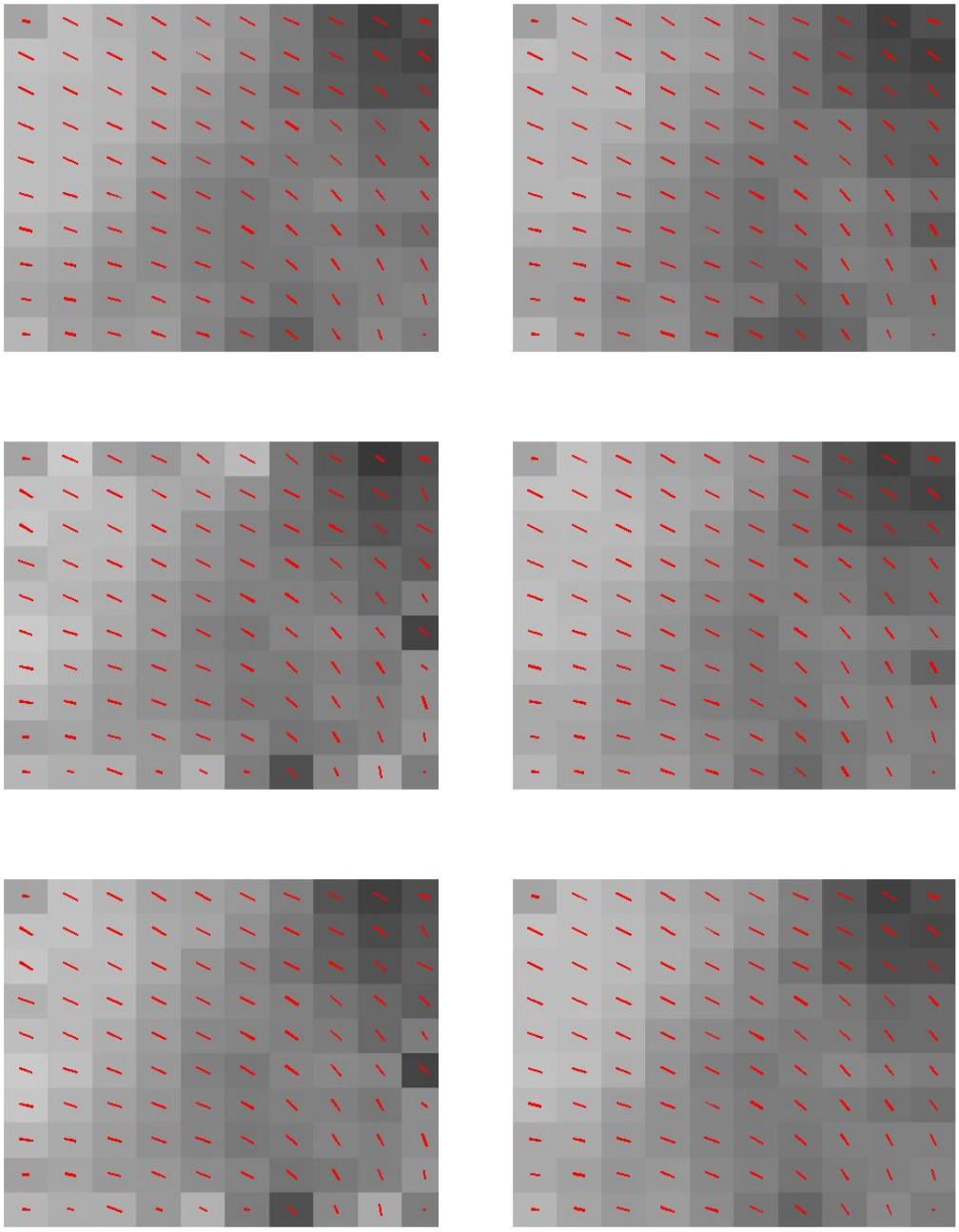


Figure 4.3 FA and v1 maps of median tensors using Euclidean(top left), power Euclidean(top right), log-Euclidean(middle left), root Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods.

The line plots of the determinant of tensors are shown in Figure 4.4. Each line represents the change of determinate of ten tensors at each row of the tensor space. The lines of Euclidean, root Euclidean and Procrustes method are smoother. In contrast, the bottom two lines of log-Euclidean and Riemannian are more twisty.

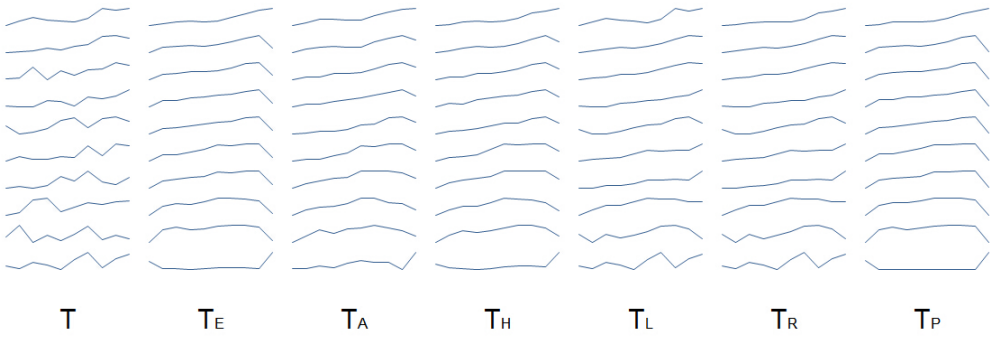


Figure 4.4 The determinant of absolute values of tensors on each row.

4.5.3 Anisotropy study

Anisotropy indices are useful to capture the diffusion anisotropy quantitatively in DTI. A most common anisotropy index is the fractional anisotropy. The range of FA is [0,1], with 0 (full isotropy) to 1 (complete anisotropy). The *Geodesic anisotropy* (GA) of a tensor **D** is the Riemannian metrics of **D** and its adjusted identity matrix I [10].

$$GA(D) = d_R(D, \sqrt[3]{det(D)}I) = \sqrt{\sum_{i=1}^{3}(\log\lambda_i - \log\bar{\lambda})^2} \tag{4.41}$$

GA is positive and not in the unit range. It is more convenient to compare the all the anisotropy indices in the unit range. Thus normalisation of GA is defined as

$$\tanh(GA) = \frac{\exp(2GA) - 1}{\exp(2GA) + 1} \tag{4.42}$$

Another anisotropy index is Procrustes anisotropy (PA) [113]. It is the full Procrustes shape metric between tensor **D** with its identity matrix, normalised with a factor $\frac{3}{2}$. PA is also in the unit range with 0 (full isotropy) to 1 (complete anisotropy)

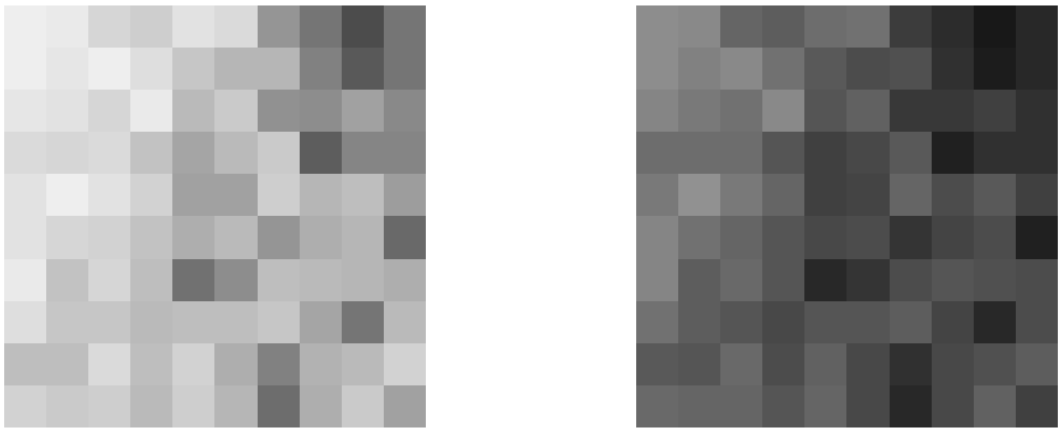


Figure 4.5 Simulated 10×10 tensor field and its FA and v1 map.

$$PA(D) = \sqrt{\frac{3}{2}} d_S(D, \frac{I}{\sqrt{3}}) \tag{4.43}$$

$$= \sqrt{\frac{3 \sum_{i=1}^{3} (\sqrt{\lambda_i} - \sqrt{\bar{\lambda}})^2}{2 \sum_{i=1}^{3} \lambda_i}} \tag{4.44}$$

We produce the tanh(GA) and PA map of the tensor field in Figure 4.5. Generally, the PA provides a darker colour map and tanh(GA) provides a lighter colour map than FA.

4.5.4 Smoothing with weighted median tensors

Now we smooth the diffusion tensor space with weighted median tensors. The same 10×10 tensor space in Figure 4.2 is used in our experiments. It has four distinct tensor at four corners and a Gaussian noise of 0.5. In the last section, we calculated the median tensor estimator using its neighbour four voxels and itself. In this section, we use its neighbouring eight voxels and itself to estimate the weighted median tensors. Then the estimators replace their voxels. The results of FA and main diffusion direction v_1 maps are generated in Figure 4.6, 4.7. The inverse distances weights are used in Figure 4.6 and Exponential weights are used in Figure 4.7.

Generally speaking, the weighted median tensors obtain much smoother images compared to the original simulated tensors in Figure 4.2. The results of inverse distance weights are smoother than the results of exponential weights. The transition of colours in FA maps is more gradual in Figure 4.6. This is the same conclusion as we draw earlier, the exponential distances function gives more weights when the distances are very close. In other words, it tends to retain its original properties. The main directions of the tensors are all moved into orders in both figures. The discrete Procrustes methods give a smoother result than the others.

4.5.5 Interpolation with equal weights

The median tensors can also be used to interpolate the tensor field. The simplest interpolation uses equal weights. We still use the same tensor field in Figure 4.2. Imaging the tensor field expands, one more row and two column are inserted between each two rows and two columns of voxels, as shown in Figure 4.8. At each 3×3 voxels (the red box), we use four voxels are used to interpolate the central voxels and the neighbouring two tensors to interpolate the other four tensors at the central voxel at edges. The results are shown in Figure 4.10. Figure 4.9 shows a interpolated result by log-Euclidean median tensors. Generally speaking, the interpolation of tensor field gives more definition of the tensor behaviours. Their transitions are more detailed.

4.6 Real data analysis

4.6.1 Smoothing with non-weighted median tensors

This study is based on the diffusion MR images from a healthy human brain. The median non-Euclidean methods are used to compute the tensor field. The region of the corpus callosum (CC) is enlarged from axial, sagittal and coronal views respectively in Figures 4.11, 4.13 and 4.15.

In general, the colour of FA and the direction of v_1 (the principal diffusion direction) of the results illustrate brain tissues more clearly. Since there are voxels of contrasting direction in Figure 4.11, the power and root Euclidean methods show a more clear pincer structure whilst others retain some dark points at the left side of CC region in Figure 4.12.

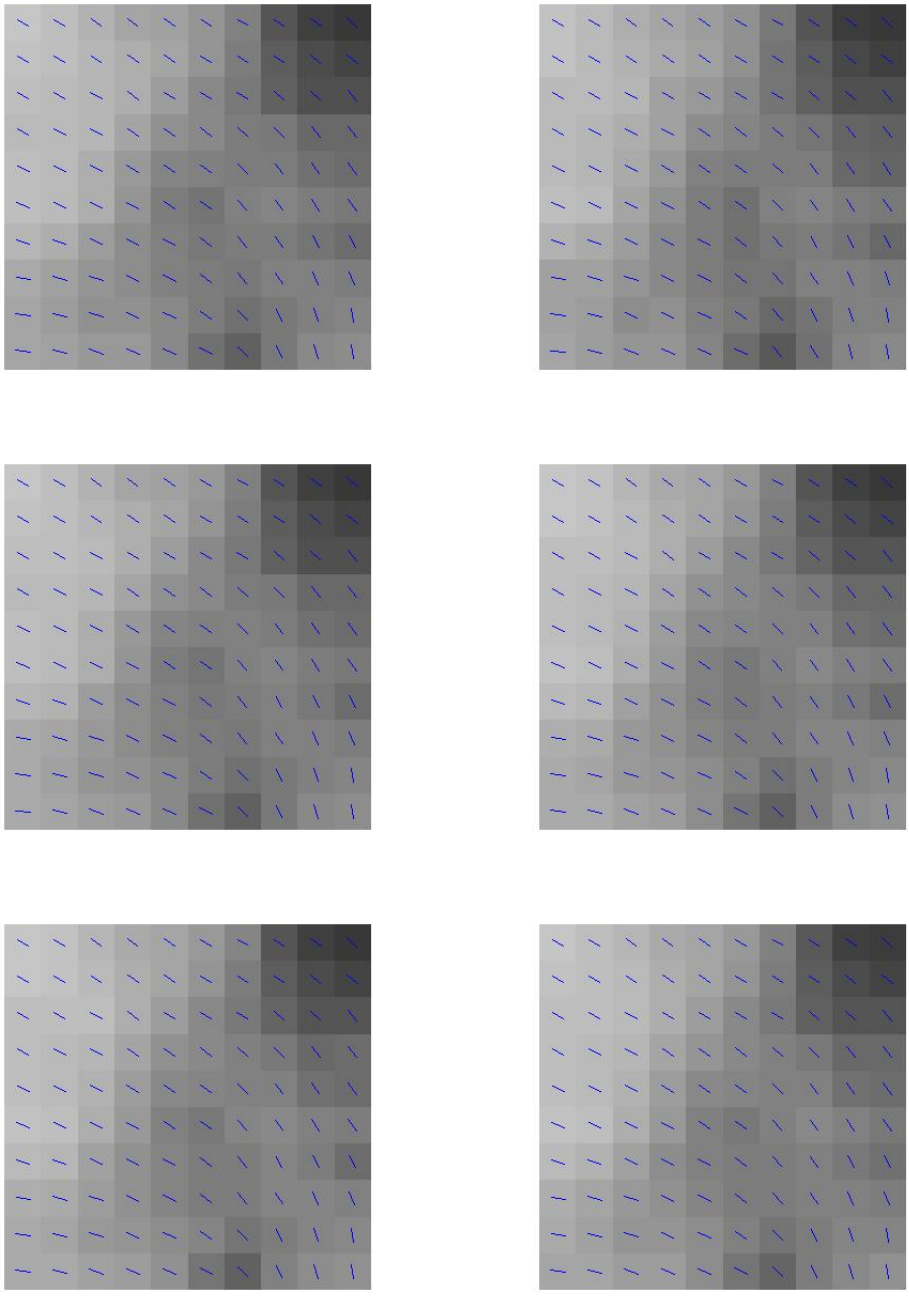


Figure 4.6 FA and v_1 maps of median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods with inverse distances weights.

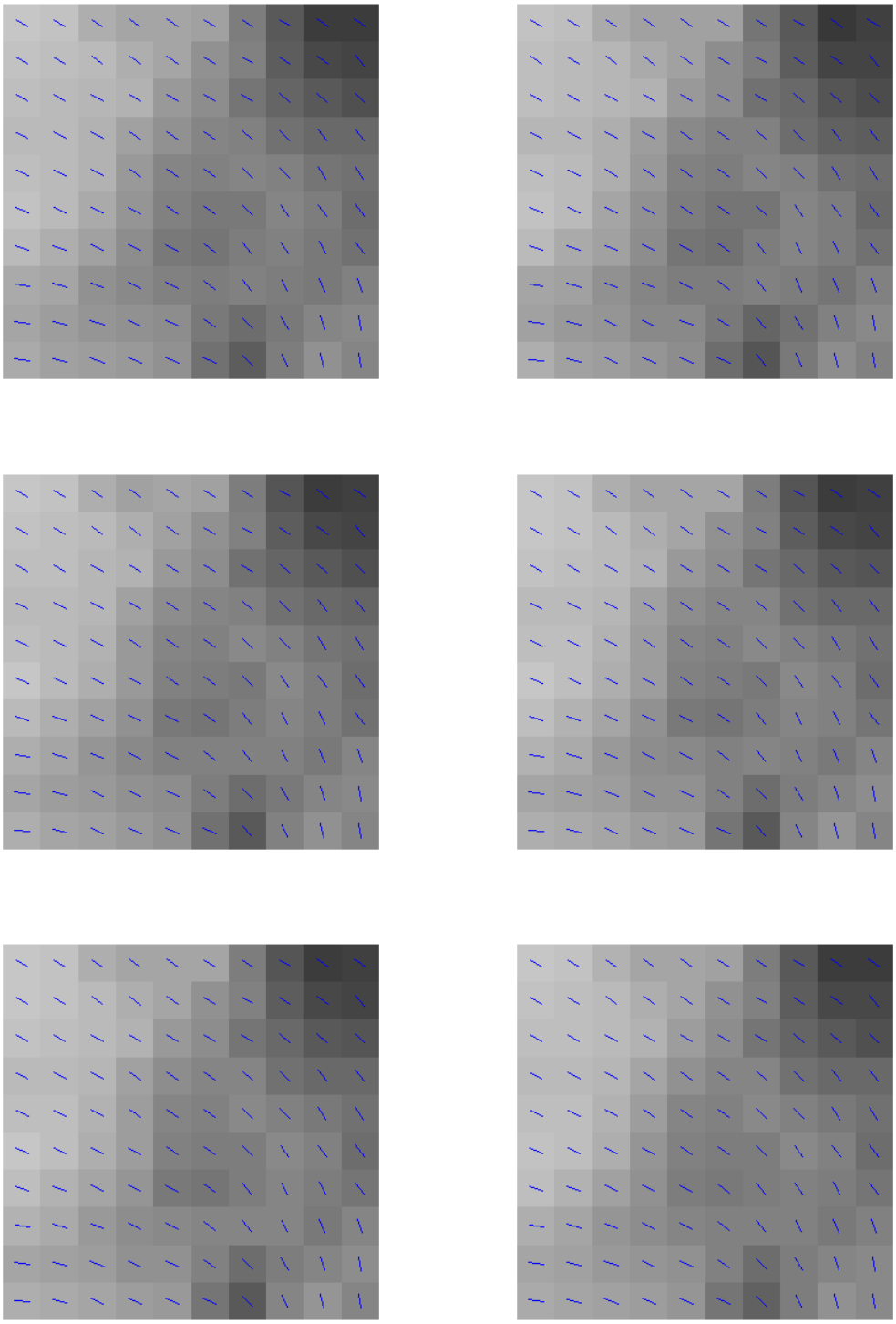


Figure 4.7 FA and v_1 maps of median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods with exponential weights.

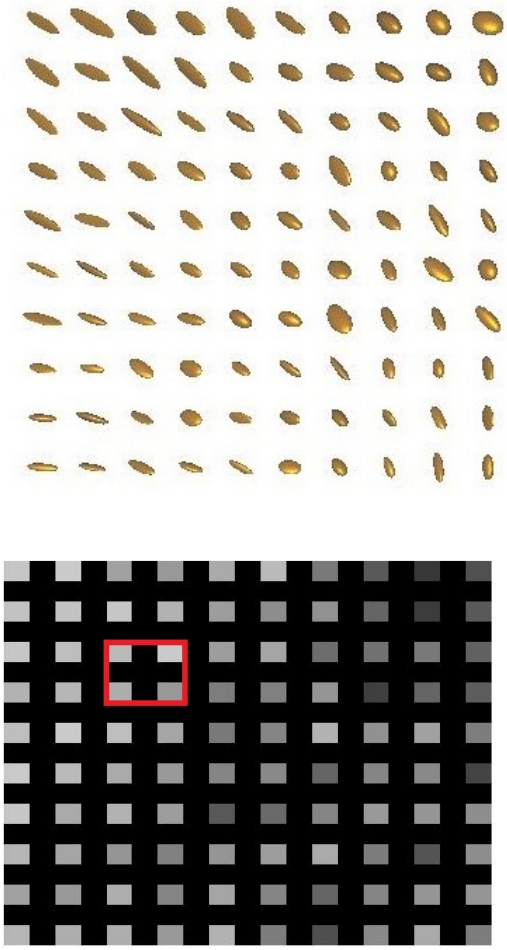


Figure 4.8 Expanded 10×10 tensor field and its FA map.

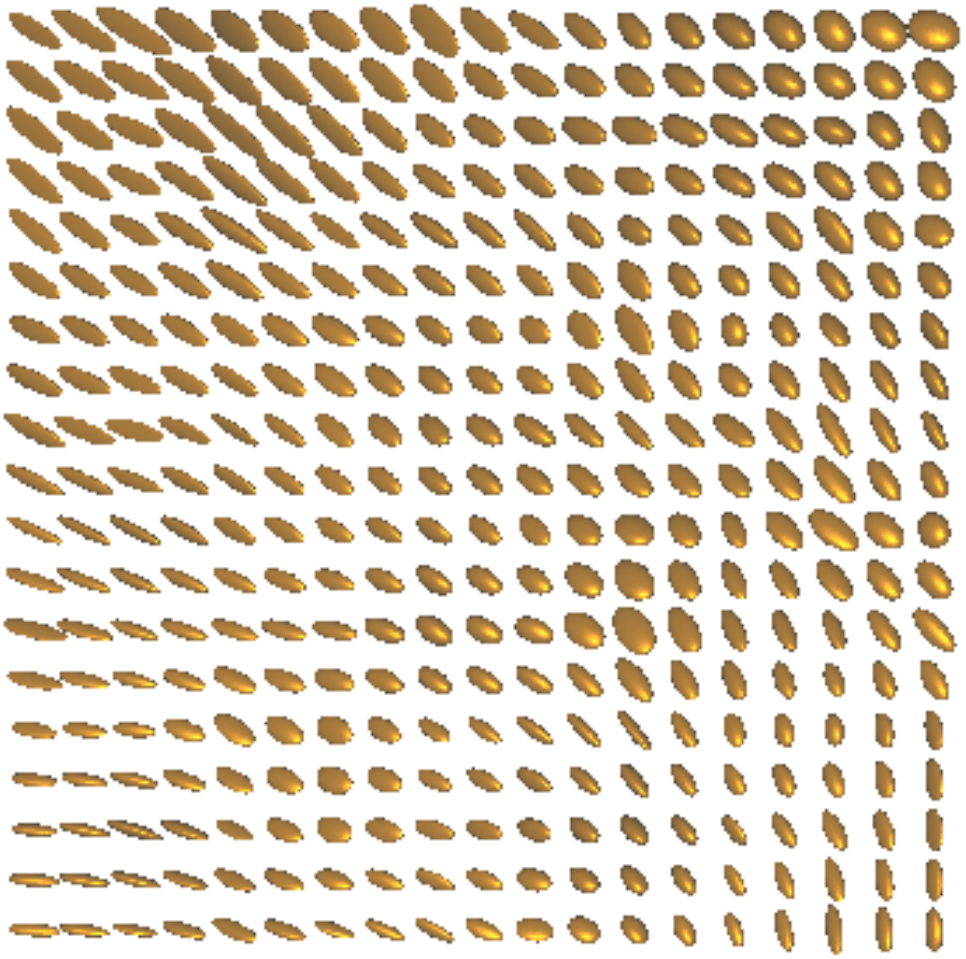


Figure 4.9 Interpolation of 10×10 tensors FA map by log-Euclidean median tensors.

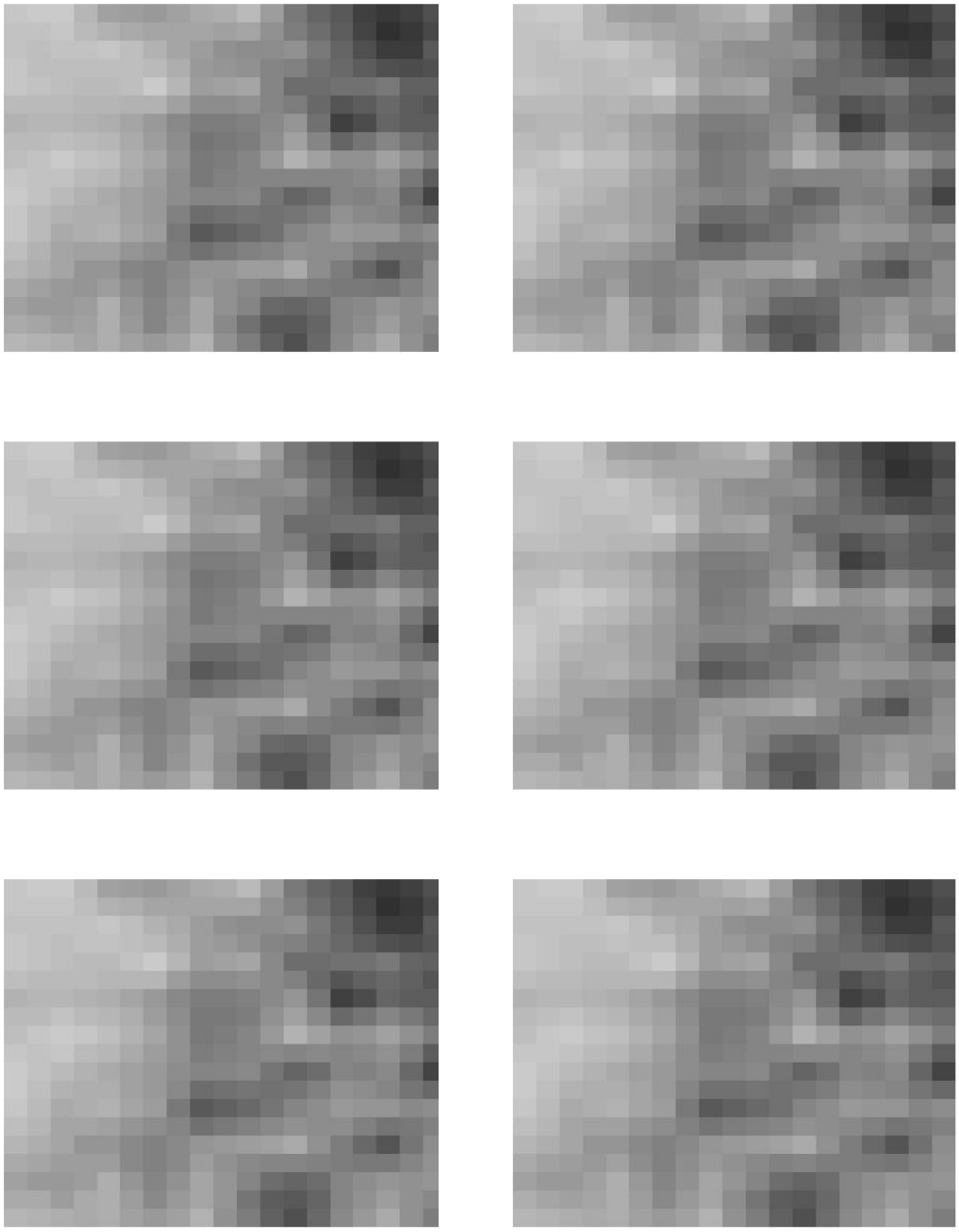


Figure 4.10 FA maps of interpolation with median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods.

The gap between cingulum and corpus callosum is clearly shown in Figure 4.14, whilst in the original DTI image (Figure 4.13), this gap is poorly defined. In Figure 4.16, it is shown more clearly that cingulum and corpus callosum are not attached, whereas, in Figure 4.15, the feature is folded.

4.6.2 Smoothing with weighted median tensors

In this section, we exam the effect of smoothing the DTI brain image by weighted median tensors in the real applications. A DTI image from a healthy human brain is used in the experiments. Three fractions of DTI images from axial, sagittal and coronal views are shown in Figures 4.11, 4.13, and 4.15. The two weighted functions, inverse distances and exponential distances, are applied in the smoothing.

From the axial view, the smoothing results from inverse distance weighting are shown in Figure 4.19 and the results from exponential distances are shown in Figure 4.21. We can see that the area of intersection of fibres (part A in Figure 4.11), is very unclearly defined. While in Figure 4.19 and 4.21, all the images are described more details in the area A. There are two discrete methods, Euclidean and Procrustes median tensors smoothing. Their results are shown in the top left and bottom right of Figure 4.19 and 4.21. As we can see, the discrete methods have poorer performances than the other non-Euclidean methods in area A and B. The results from power Euclidean are slightly blurred in area A. The results from root Euclidean, log-Euclidean and Riemannian median tensor clearly describe the intersection of fibres.

Figure 4.18 shows the tensors in 3D of the human brain DTI image in Figure 4.11. Each ellipsoid tensor describes the behaviour of the water molecule of each voxel. It is noticeable that some abnormal tensors behave extremely different to their surroundings. These incorrect tensors influence their FA and v_1 maps so that we can not see the insight structure of human brain clearly enough. Two smoothing results are illustrated in Figure 4.23. The top graph is the result from Euclidean distance method, corresponding to the top left graph in Figure 4.19. The tensor graph is more harmonious. All the abnormal tensors are well behaved expect in area A. The bottom graph gives the result from root Euclidean methods. The transaction of tensors is smooth. It is easy to see the movement of the water modules.

The FA values of each smoothing results from axial view are also shown in Figure 4.20 and 4.22. In Figure 4.11, the majority of FA values of original DTI image are from 0 to 0.2. In comparison, the FA values of smoothing are still much smaller than 0.2, but the number of FA values between 0 and 0.1 is less and the number of FA values between 0.4 and 0.6 is

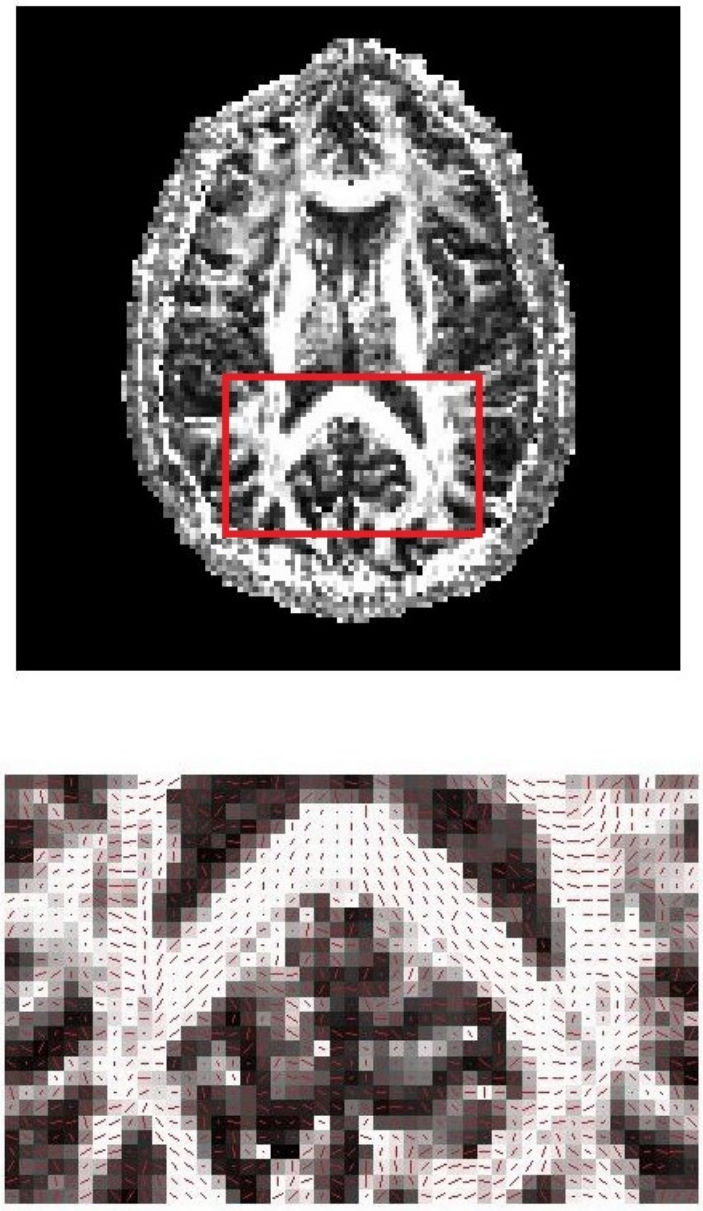


Figure 4.11 Real DTI images (axial view) and its enlarged splenium of corpus callosum region.

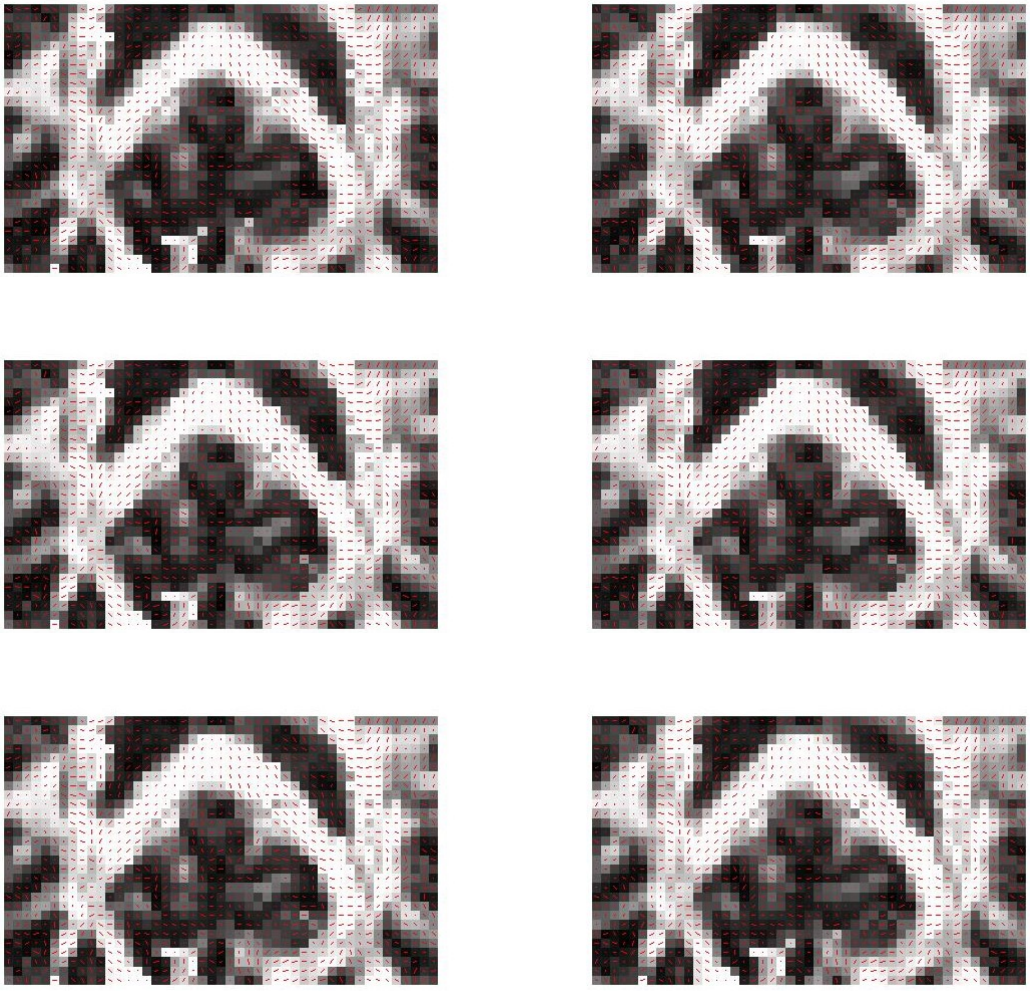


Figure 4.12 FA and v_1 maps of median tensors using Euclidean (top left), power Euclidean (top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian (bottom left) and discrete Procrustes (bottom right) distance methods.

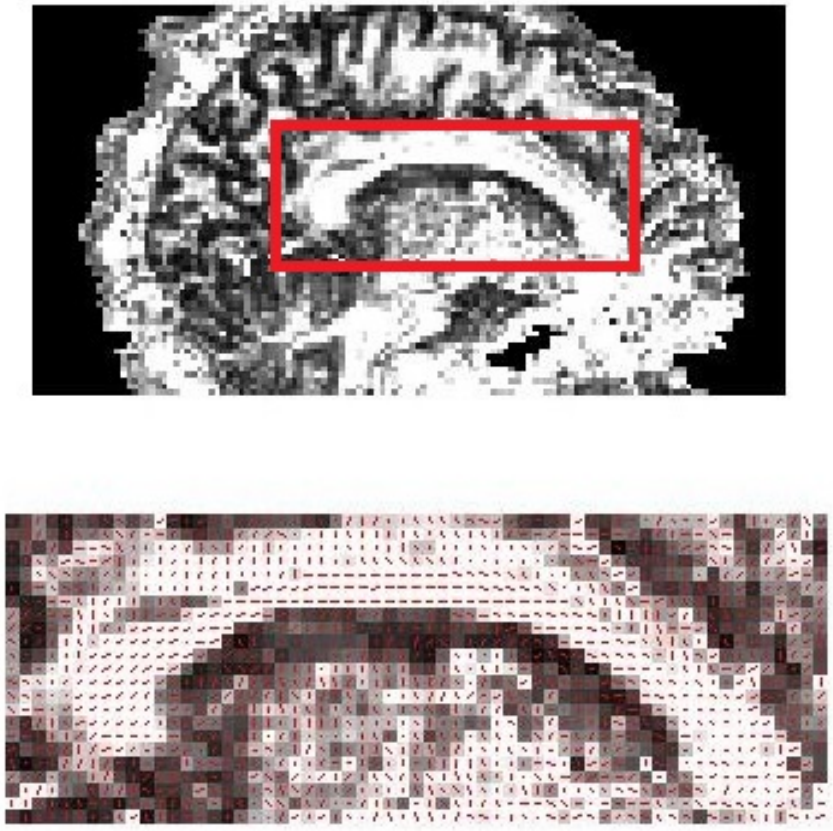


Figure 4.13 A real DTI images (sagittal view) and its enlarged corpus callosum region.

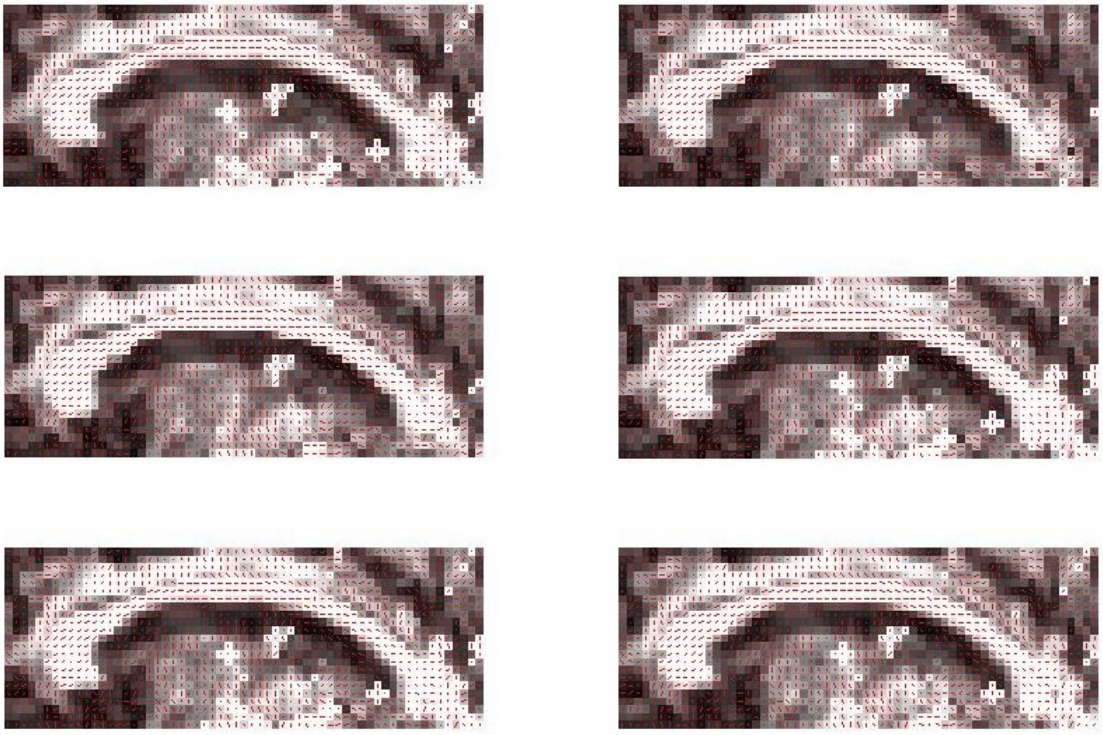


Figure 4.14 FA and v_1 maps of median tensors using Euclidean (top left), power Euclidean (top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian (bottom left) and discrete Procrustes (bottom right) distance methods.

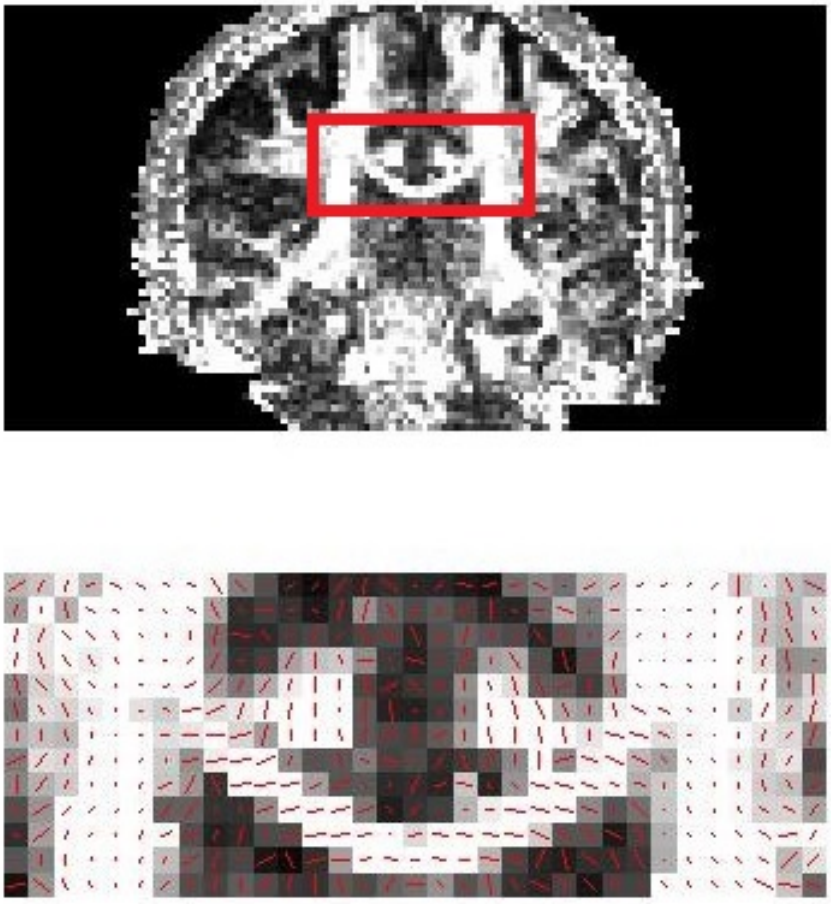


Figure 4.15 A real DTI images (coronal view) and its enlarged corpus callosum region.

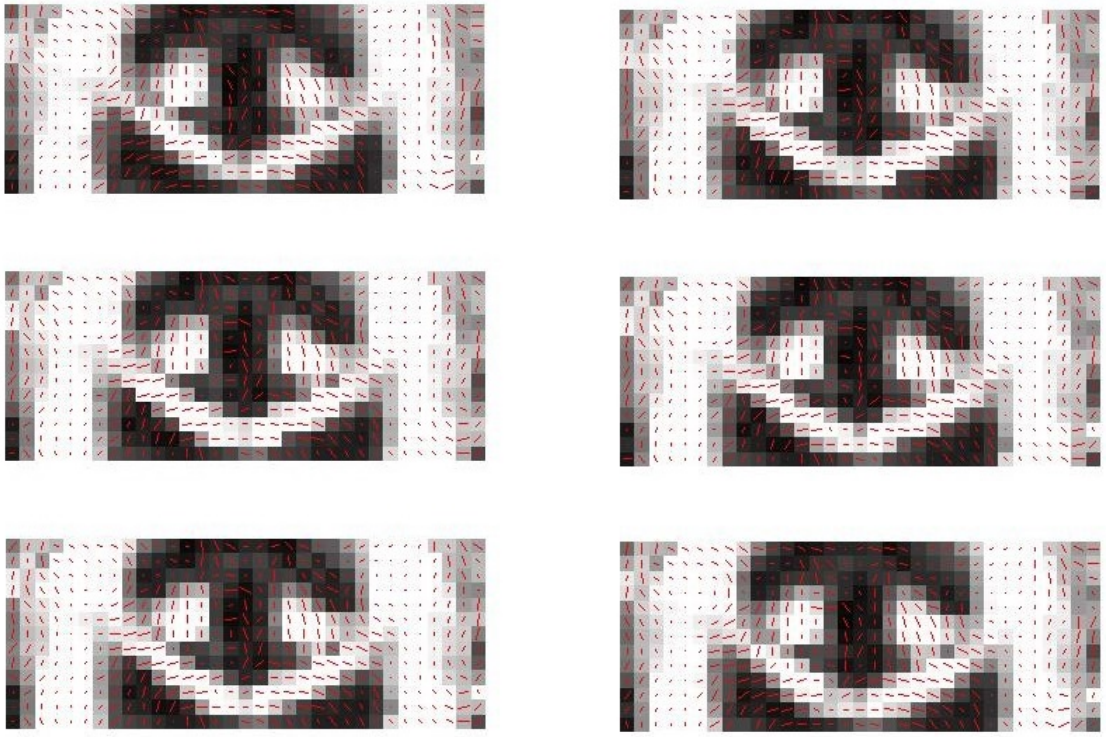


Figure 4.16 FA and v_1 maps of median tensors using Euclidean (top left), power Euclidean (top right), root Euclidean (middle left), log-Euclidean (middle right), Riemannian (bottom left) and discrete Procrustes (bottom right) distance methods.

greater. Comparing the results of two weighted functions, their FA and v_1 maps are similar in terms of each distance methods. The patterns of the histogram of FA values are also similar. The FA from exponential distance weighting is slightly greater the FA from inverse distance weighting functions.

Similarly, a segment of original DTI brain image from the coronal view is shown in Figure 4.24, and the two smoothed results calculated from the inverse distance function are shown in Figure 4.25. The smoothed images apparently better illustrate the different tissues inside the brain. In the original graph, although area C and D are both white in FA maps, the directions of v_1 are completely different in these two regions. The layer between C and D is poorly defined. In contrast, the smoothing images show clearly that there exists a layer between cingulate gyrus (area C) and corpus callosum (area D) and they are not connected. The colour of this layer is darker in Euclidean smoothed result than log-Euclidean smoothed result in Figure 4.25. There is another obvious difference in area E between these two smoothed images. The Euclidean result shows a large contrast in FA values in E and its surroundings, while the log-Euclidean result gives a better-smoothed image in terms of both FA and the directions of v_1.

Another fraction of human brain DTI image from sagittal view and its smoothed result by the Riemannian median tensors method are shown in Figure 4.26. The directions of the tensors in cingulum (area F) and genu of corpus callosum (area G) are different. The original DTI images indicate there is a layer between these two regions but not fully demonstrated. The smoothed result from the Riemannian median tensor method is better illustrated the inner brain structure. It is easy to tell area F and G are not connected in the bottom graph in Figure 4.26.

Generally speaking, the images become more contrasted and the shapes of brain tissues are clearer. In Figure 4.17, the discrete Procrustes method shows a more clear U-structure while others remain some unconscious points at the left side. The reason is that it may involve a totally different direction tensors when using its surrounding tensors to calculate the median tensors at the edge between two different tissues.

4.6.3 Interpolation

The non-Euclidean median tensors method can also use to interpolate the real brain DTI images to give a higher resolution. We use median tensors to construct new diffusion tensors based on the original data. In Figure 4.27 and Figure 4.28, the cingulum (cg) is clearly distinct from the corpus callosum (cc) in the interpolated map than its original map. It proves

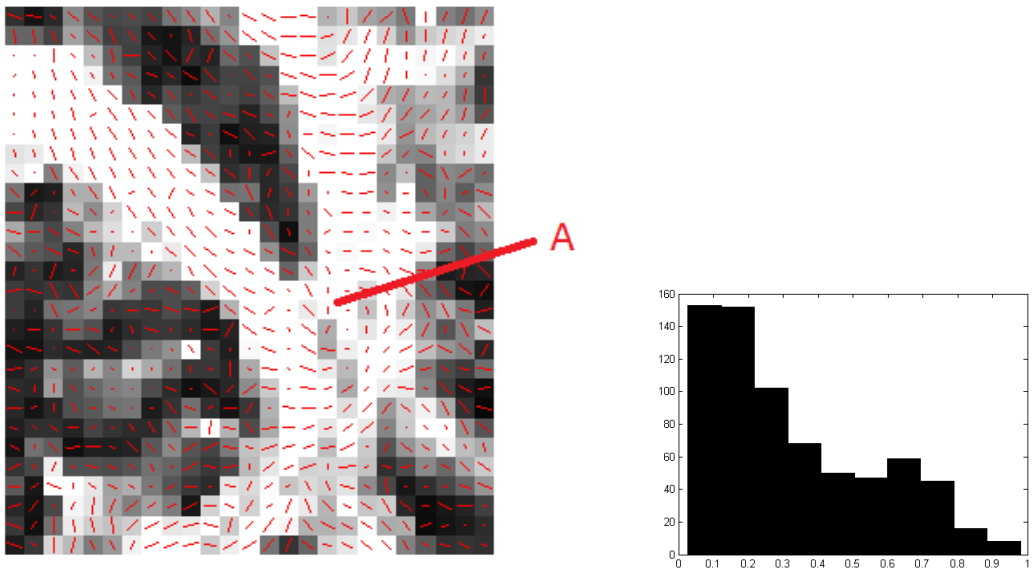


Figure 4.17 FA and v_1 maps of human brain (left) and the histogram of its FA values (right).

that there is a layer between these two different tissues. It is more clearly to see it on the FA map of non-Euclidean median estimators than the interpolation map. In Figure 4.28, several abnormal dark points are occur at the white tissues in the map of power Euclidean metric. Thus other methods are more recommended in the interpolation of DTI images.

4.6.4 Validation

In this section, we perform two experiments to validate various Euclidean mean and non-Euclidean median tensor estimators. In the first study, we select voxels in the CC region manually to analyse the effectiveness of our methods. Table 4.10 shows the RMSE of FA, determinant and orientations of estimators in the selected CC region shown in Figure 4.29. The RMSE of orientation (ϕ) is given by

$$RMSE(\phi) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \|\phi_i - \phi\|^2} \tag{4.45}$$

ϕ_i is the smallest angle between orientation of $\hat{T}_i$ and a reference orientation.

As the voxels we selected are in the corpus callosum, their estimated tensors should have similar properties, i.e. large FA, similar size and pointing at the similar directions. Thus they

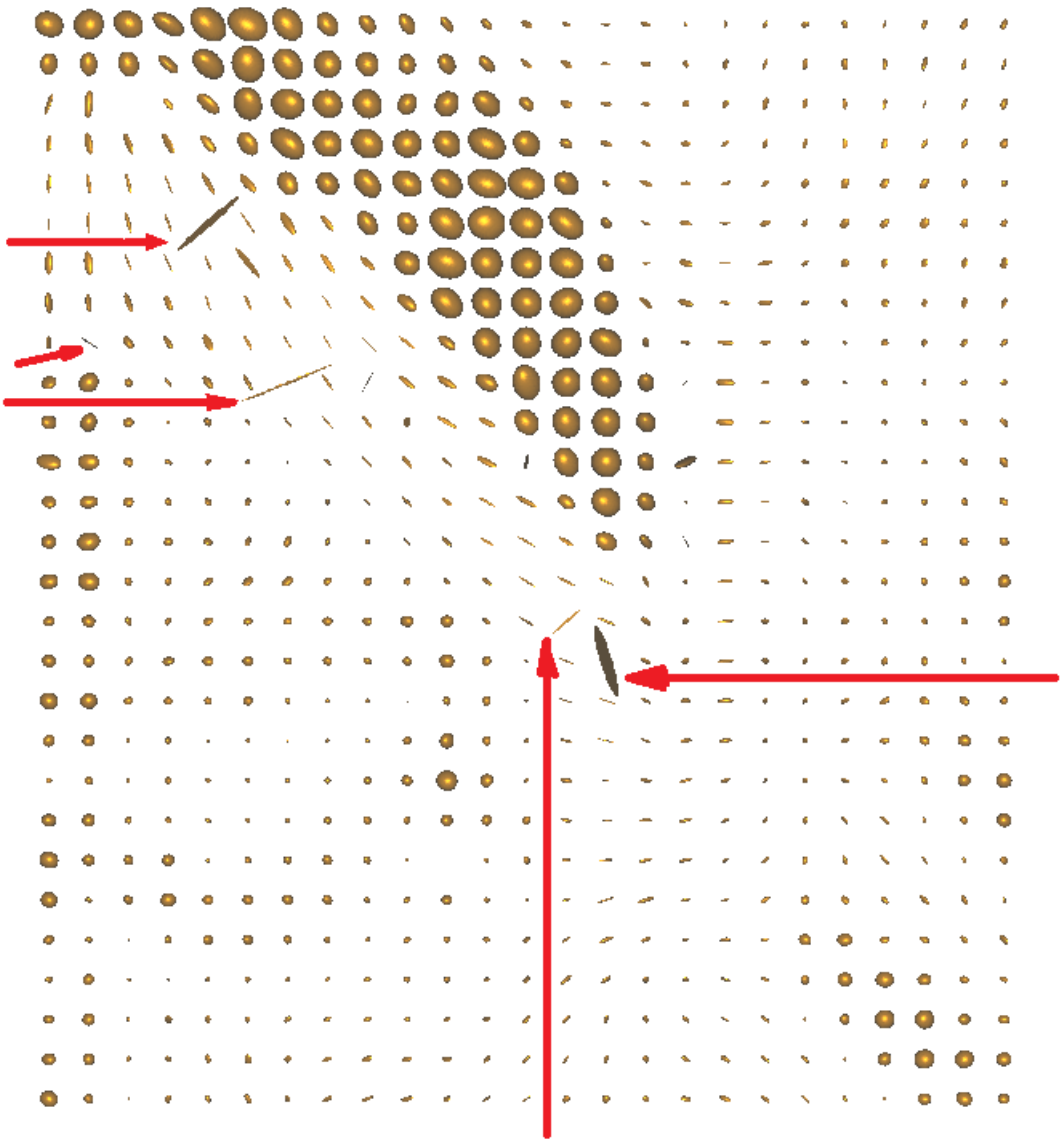


Figure 4.18 The tensor field of a human brain.

Figure 4.19 FA and v_1 maps of smoothing with inverse distance median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods.

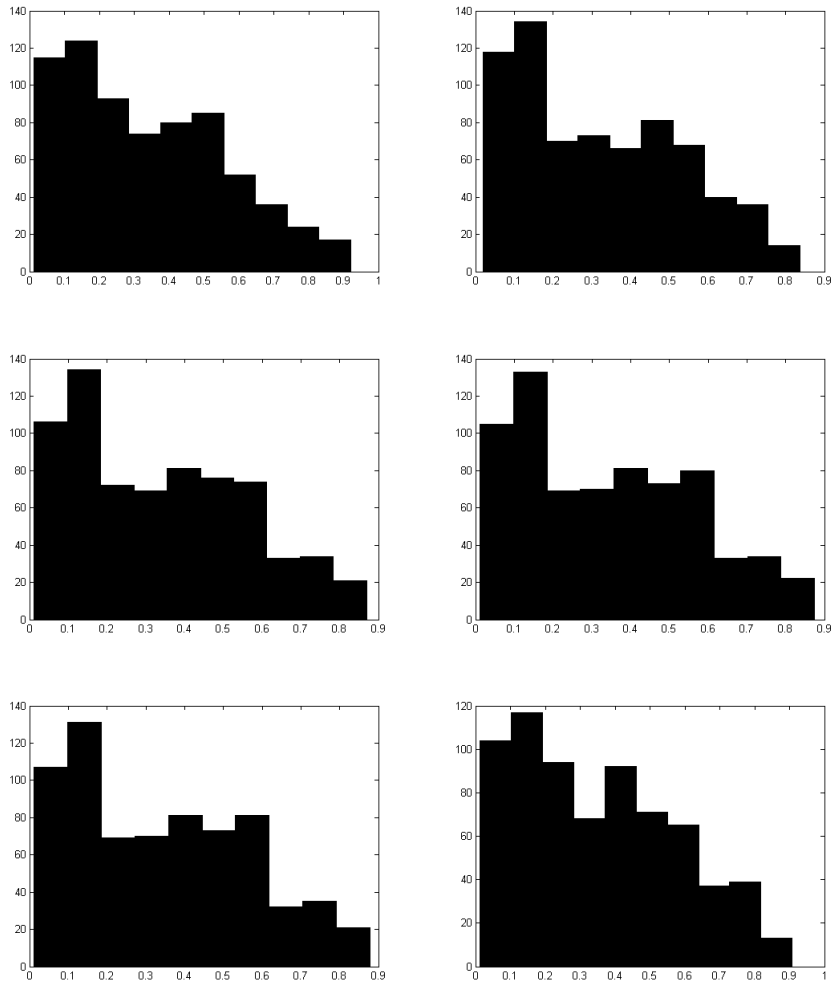


Figure 4.20 Histogram of FA of smoothing with inverse distance median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods.

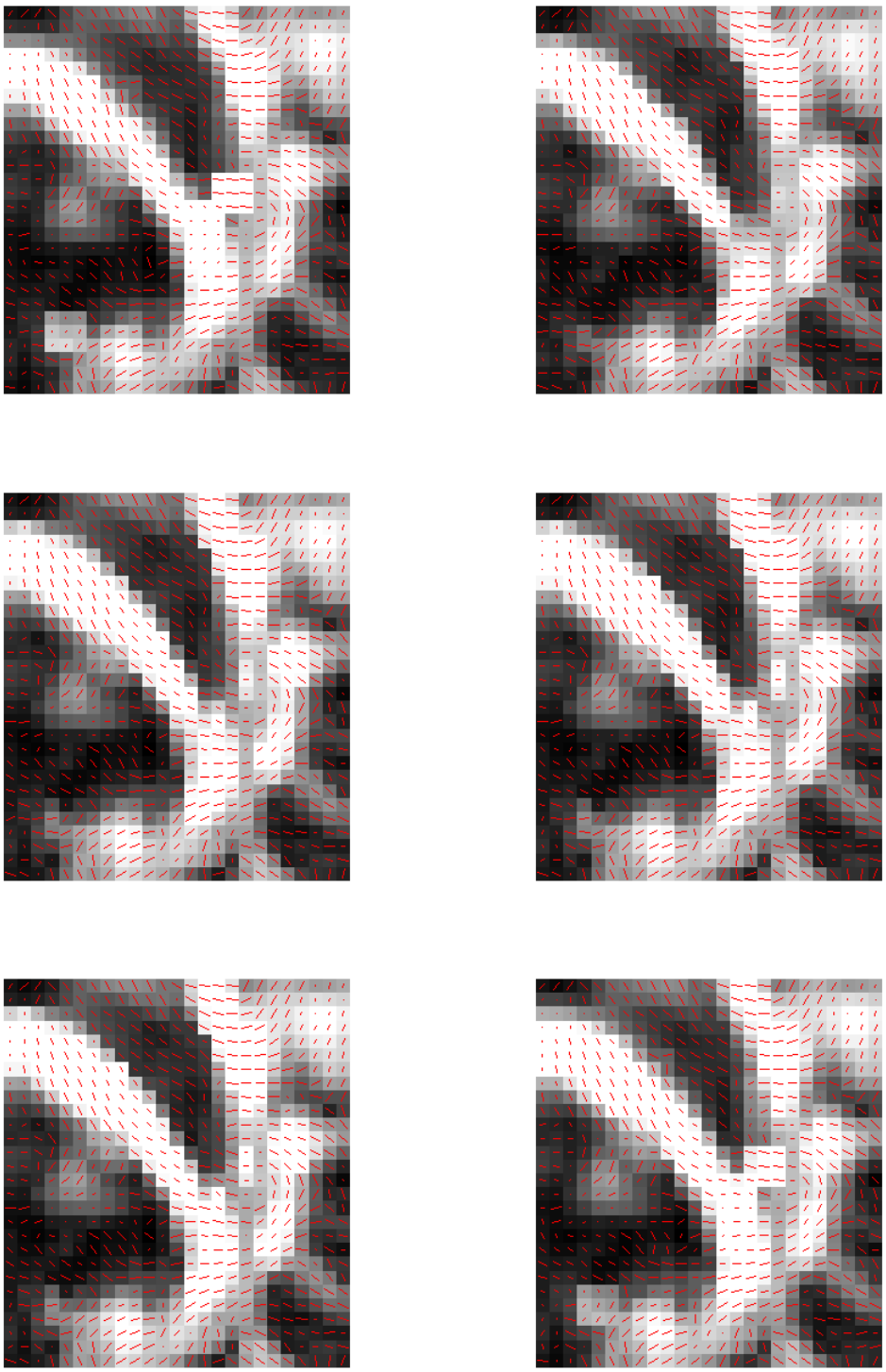


Figure 4.21 FA and v_1 maps of smoothing with exponential weighted median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods.

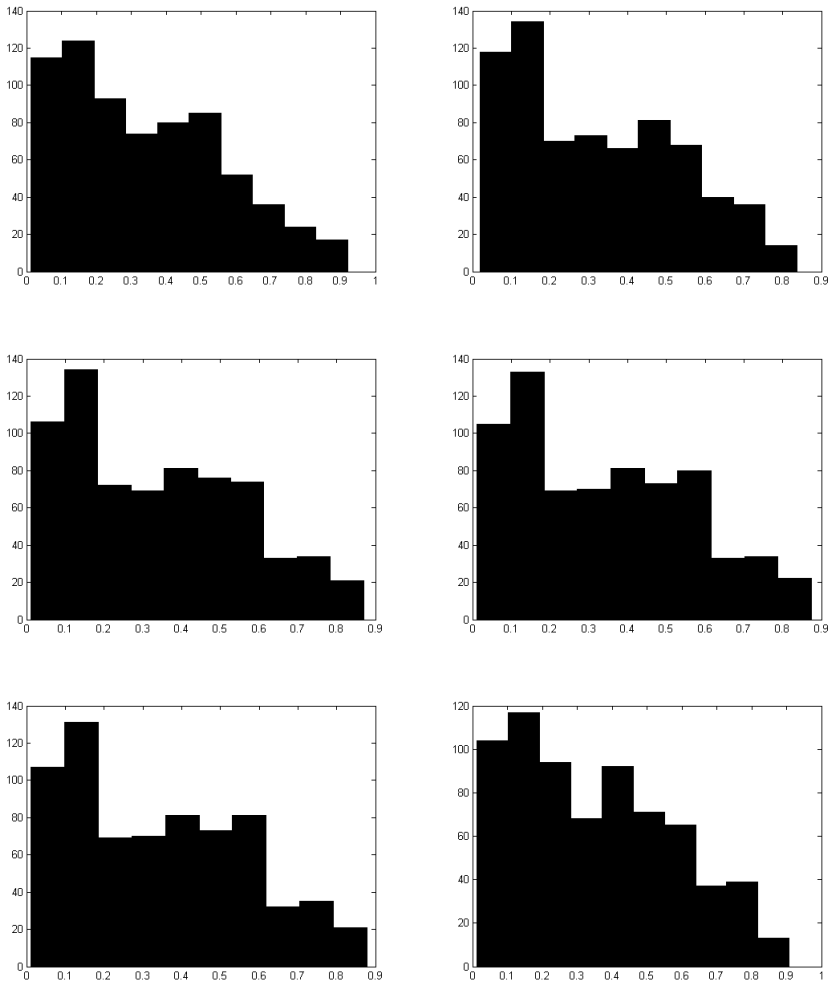


Figure 4.22 Histogram of FA of smoothing with exponential weighted median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods.

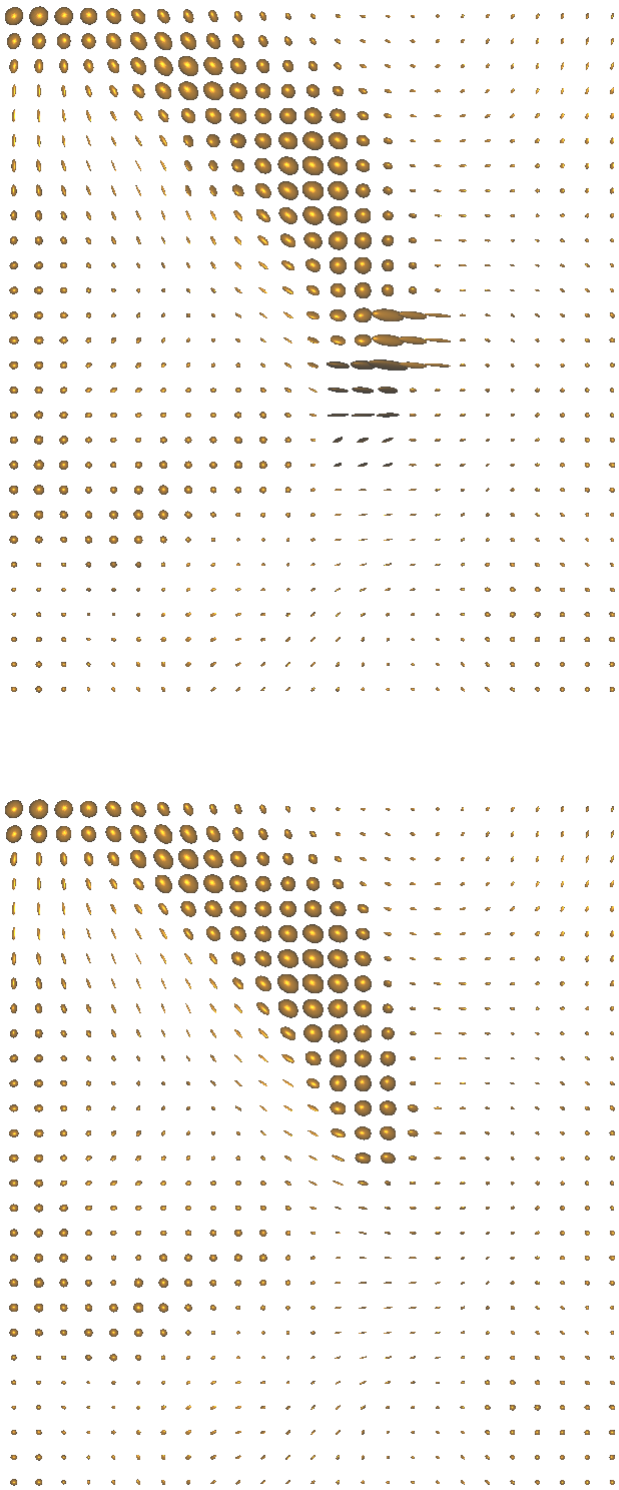


Figure 4.23 Smoothed tensors from Euclidean (top) and square root (bottom) median tensors methods.

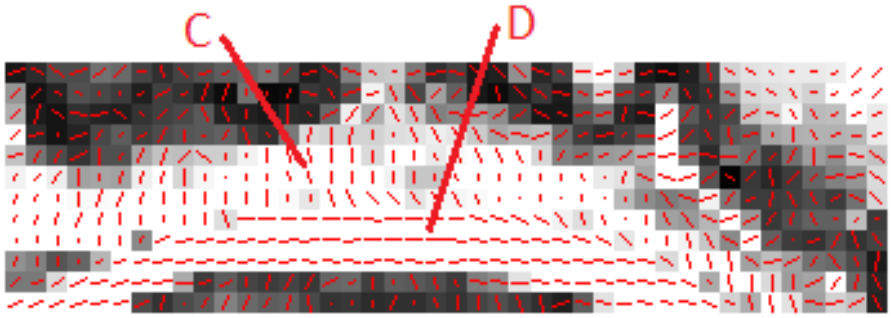


Figure 4.24 FA and v_1 map of the human brain.

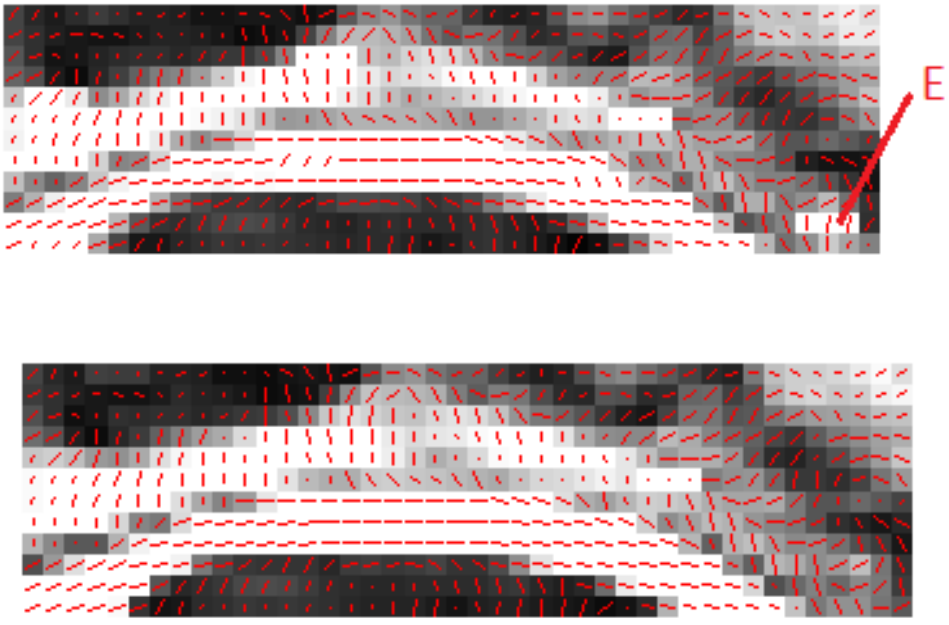


Figure 4.25 FA and v_1 map of smoothing with inverse distance median tensors using Euclidean (top) and log-Euclidean (bottom) methods.

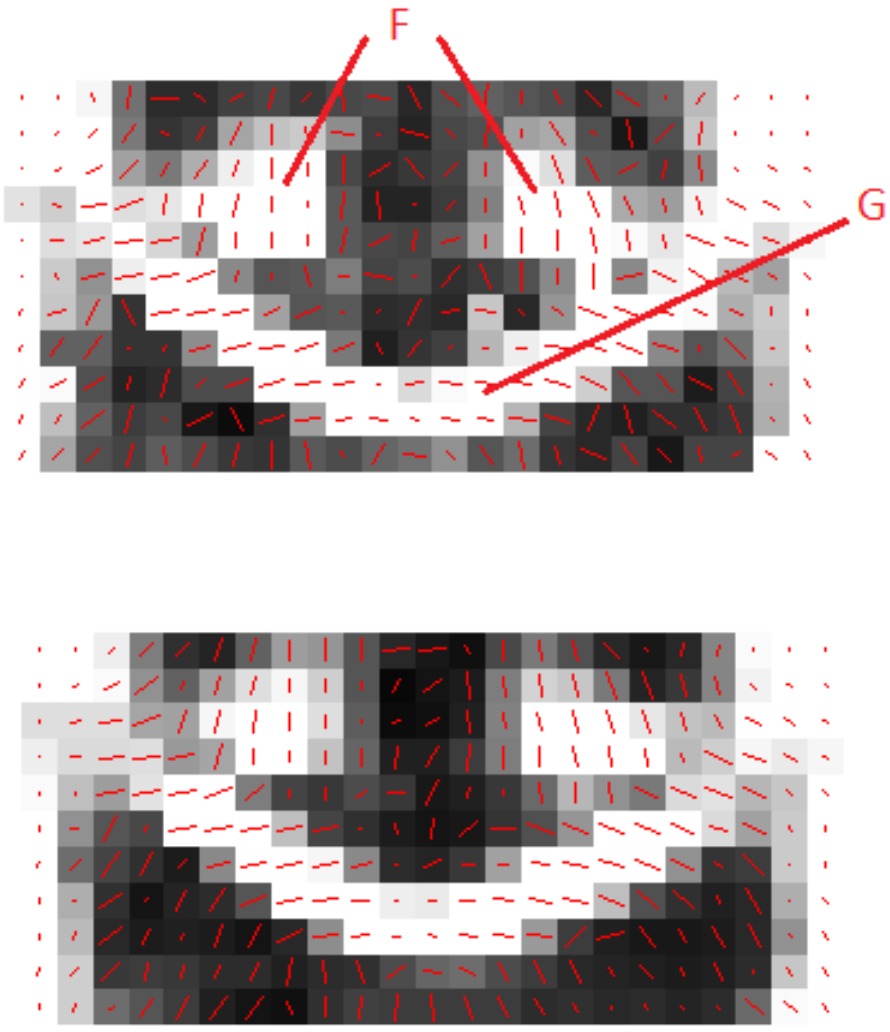


Figure 4.26 FA and v_1 map human brain (top) and its smoothed image by Riemannian median tensors method (bottom).

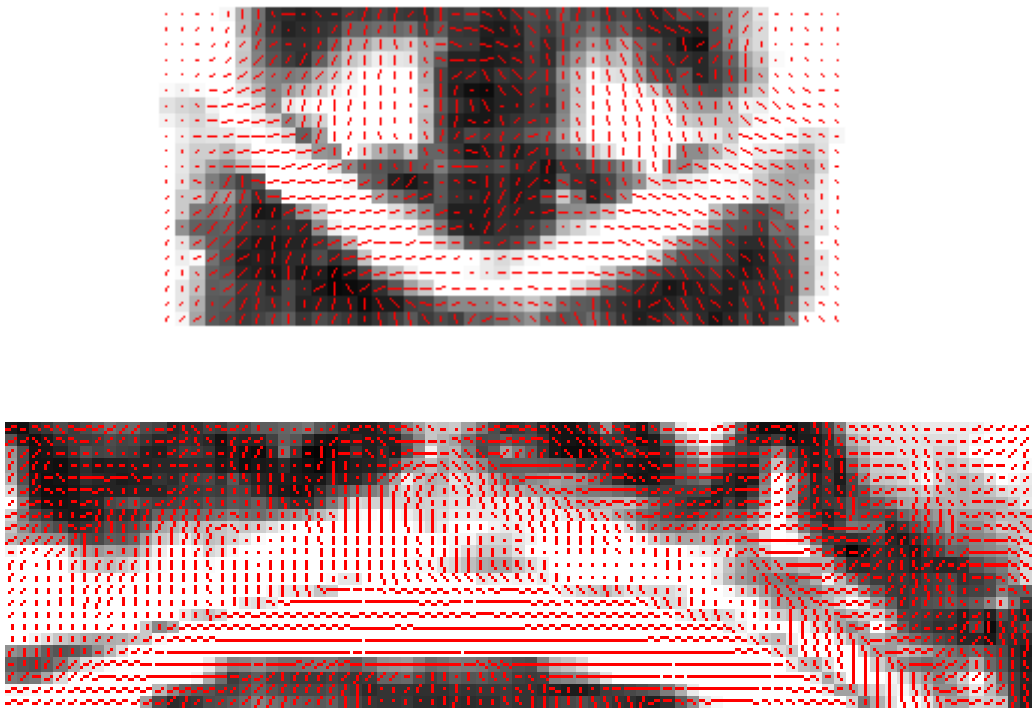


Figure 4.27 FA and v_1 map of interpolation with median tensors using root Euclidean distance methods.

Figure 4.28 FA and v_1 maps of interpolation with median tensors using Euclidean(top left), power Euclidean(top right), root Euclidean(middle left), log-Euclidean(middle right), Riemannian(bottom left) and discrete Procrustes(bottom right) distance methods.

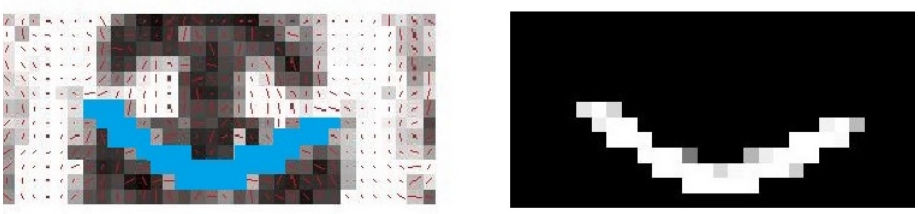


Figure 4.29 The selected voxels of corpus callosum.

$\sigma = 5000$	FA	Det(T)$\times 10^{-11}$	Orientations
T_E	0.0203	17.945	3.7006
T_A	**0.0165**	9.5933	3.8193
T_H	0.0171	**7.7034**	**3.4810**
T_L	0.0180	13.826	3.9222
T_R	0.0180	15.537	3.7929
T_P	0.0193	11.064	3.5915

Table 4.10 Root mean square error of fractional anisotropy, determinant of tensor and angle of principal directions of Euclidean mean, power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes median tensors tensors.

should have small RMSE. The less RMSE, the better estimator. Table 4.10 describes that the power Euclidean median tensor estimator has the least root mean square error in FA. In terms of determinant and orientations, root Euclidean median tensor estimator has the least RMSE. These two methods both give the least two figures in terms of FA and determinant. In terms of orientations, the result of power Euclidean method is not desirable among the six methods, whilst the result of root Euclidean is the best. Thus we conclude that the root Euclidean metric is the best recommended in tensor field processing. The power Euclidean metric is also recommended when producing FA maps.

The second experiment performed is to validate the robustness of different Euclidean mean and non-Euclidean median tensors in tensor field processing. The ROI is the part of CC region and its surrounding voxels shown in Figure 4.30. The tensors in the ROI can be classified into two groups. The tensors at the CC region have relative high FA value and the surrounding tensors have relative low FA value. The MD of tensors at CC region is around two or three times greater than MD of surrounding tensors. The orientations of the tensors in these two groups are also quite different. We add Gaussian noise to the raw data with standard deviation $\sigma = 5,000$. The Cholesky non-linear least squares method is used to estimate the tensors to ensure the positive semi-definiteness of diffusion tensors, then various Euclidean

Figure 4.30 The corpus callosum region and its surrounding of a human brain.

and non-Euclidean techniques are used to smooth the noisy tensor field with exponential weights ($A = 3, B = 0.5$). We compare the results with smoothing the raw tensor field. The FA, MD and orientations of the estimated median tensors are calculated. Table 4.11 shows the mean of absolute value of differences between the median tensors estimated from noisy data and raw data. The power Euclidean median tensor estimator has the least difference in FA. The root Euclidean median tensor estimator has the least differences in MD, and power Euclidean, log-Euclidean and Riemannian median tensor estimators also have satisfied results in MD. The Procrustes median tensor estimator has the least difference in terms of orientations, followed by power and root Euclidean estimators. Overall, the power Euclidean, root Euclidean and Procrustes median tensor estimators each have advantages in processing tensor field. However, the Procrustes median tensor estimator produces a large FA difference between noisy and raw data. The power and root Euclidean median tensor estimators have reasonably good results in terms of FA, MD and orientations, thus these two methods are preferable.

In this section, two applicants studies are performed to validate various Euclidean and non-Euclidean techniques. In conclusion, the power and root Euclidean median tensors have the least differences between the noisy and raw data. These two methods have higher robustness compared to other methods. The power Euclidean median tensor is recommended especially when producing FA map.

$\sigma = 5000$	FA	MD	Orientations
T_E	0.2026	0.0941	21.8888
T_A	**0.1456**	0.0925	19.4568
T_H	0.1552	**0.0923**	19.4592
T_L	0.1602	0.0924	19.7777
T_R	0.1607	0.0924	19.7492
T_P	0.1777	0.0996	**18.4097**

Table 4.11 The mean of absolute value of difference in terms of FA, MD and orientations between tensors estimated by Euclidean mean, power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes median tensors from raw data and noisy data with noise standard deviation $\sigma = 5,000$.

4.7 Summary

In Chapter 4, we are focused on tensor field processing. First, we define the median non-Euclidean covariance matrix. The estimation techniques include power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes metrics. The tensor estimators are parameterised when necessary to ensure the symmetric semi-positive definiteness property. The weights, inverse Euclidean distances weighting and exponential weighting functions are also considered. The geodesic anisotropy and procrustes anisotropy have been compared with fractional anisotropy index. The weighted non-Euclidean median tensor estimators are compared with Euclidean mean estimators in simulation studies and real data in terms of the determinant, MD, FA and main orientation. The comparison between different estimations of the median tensor is made from a single covariance tensor estimation to the tensor field processing. The non-Euclidean distances between each two estimated tensors are compared in the tables. The non-Euclidean median tensor estimators are used to smooth and interpolate the tensor field with weighting functions in simulation studies and real data. We also validate our various non-Euclidean metrics. The root-mean-square deviation of FA, determinant and orientations of tensors are compared in the corpus callosum region. Extra Gaussian noise is added to the diffusion signals. The Euclidean mean and various non-Euclidean median tensors are used to smooth the tensor field. We compare the results with smoothing original raw tensor field.

In generally, the Euclidean estimators are undesirable and problematic in the simulation and real data analysis due to a tensor swelling effect: the Euclidean averaging of tensors results in larger determinant than the original ones [5]. The non-Euclidean median tensor

estimators improve the illustration of DTI images. The log-Euclidean and Riemannian estimator produce similar results. They have an advantage of tensor field processing in terms of MD and determinant. These two estimators are more likely to be influenced by outliers. The Procrustes estimator is appropriate in a small variation data set. The power and root Euclidean estimators lead to similar results. The power Euclidean estimators have an advantage over tensor field processing in terms of FA, but they are not preferable for processing when the size and orientation are of interest. The root Euclidean median estimator is shown to be preferable overall, especially in terms of MD and orientations. It can effectively exclude outliers and clearly show the important structures of the Brain. The power Euclidean and root Euclidean techniques are recommended among six methods.

The median estimators are expected to be useful in manifold-valued data applications. The weighted median estimators achieve better results than equal weighted estimators. The data shows that the exponential weighting function is preferable especially in terms of the orientation of tensor. The improvement by weights of estimators depends on the different metrics.

Chapter 5

Conclusion and future work

5.1 Conclusion

This research is focused on tensor estimation of diffusion tensor imaging. In Chapter 1, we first introduced the background of Magnetic Resonance Imaging, then looked in detailed through Diffusion Tensor Imaging. The diffusion tensor Gaussian model and tensor eigensystem are introduced. The traditional linear and non-linear least squares methods for tensor estimation are described. Four important visualisation indices, mean diffusivity, fractional anisotropy, relative anisotropy and volume ratio are described. The fibre tractography and clinical applications are also illustrated in this chapter. The commercial and non-commercial software for DTI are summarised in tables. The cost, main functions, implemented language, input data format and operation, system requirements and main users of each DTI tools are given.

The diffusion tensor is a symmetric and positive semi-definite matrix, describing the Brownian motion of water molecules in the diffusion tensor magnetic resonance imaging (DT-MRI). However, it may be corrupted in presence of the noise arising in the measurement of diffusion attenuations. Since the commonly used least squares methods are not reliable in detecting the outliers, we propose robust methods with Forward Search algorithm to improve the tensor estimation procedure in Chapter 2. We first reviewed the noise and artifacts occurring in the DTI signal measurements. The robust statistics is briefly introduced. The robust indices and estimators are illustrated. The least median squares regression, the least trimmed squares regression and the forward search algorithm for robust estimation and its test criterion is illustrated in the methodologies. Simulation studies are carried in three cases: different fractional anisotropy, volume and orientation of the tensors. The properties of

estimators are compared with the root-mean-square errors. A series of outliers are added in the simulation. In the real data analysis, we first compare the estimators of a single diffusion tensor and read the pattern of its neighbouring tensors. Then we analyse the real data arising in the diffusion signals and conclude the appropriate outliers selection methods and threshold values. The comparison is applied to three DTI images from the coronal, sagittal and transverse view respectively. The absolute value of difference in terms of FA, MD and orientation between the various estimators are shown in tables and figures. We also validate the robust and non-robust methods in the real data by adding extra noise to the diffusion signals. The mean of absolute value of differences in FA, MD and orientations are shown between tensors estimated from noisy data and raw data. In general, the results show that the robust estimators have better results than the traditional estimators. The NLS method is problematic at estimating the determinant. It tends to have a larger volume than the real tensor. The LMS method is preferable when there are more outliers. LTS regression is recommended as it has a better estimation of determinant of tensors than the LMS method. When the noise level is small, the LLS is the best regression to estimate diffusion tensors. When the noise becomes large, robust methods are better estimations than the non-robust methods. The LTS is the most robust among four methods.

In Chapter 3, we focus on the intersection of the multiple fibres in the human brain. We have reviewed the recent techniques of multiple tensor models problems, including high angular resolution diffusion weighted imaging, orientation distribution function, mixture Gaussian model, diffusion spectrum imaging, q-Ball imaging and spherical deconvolution methods. As the information of neighbouring tensors is greatly influential in the estimation of multiple tensors in the intersection and has been ignored in the previously mentioned methods, we develop Bayesian single and double tensor models which take into account the neighbouring tensors. The background of Bayesian statistics has been introduced. In the Bayesian models, the priors use the information of neighbouring tensors, and the variance is represented by the inverse Gamma distribution. The likelihood function and posterior distribution of single and double tensor models have been deduced. Several indices of model selection criterion are stated. We also proposed Bayesian neighbour algorithm to estimate the single and double diffusion tensors with model selection available. The starting point and the phase portrait have been discussed and the procedure of the algorithm is stated in the framework. In the simulation studies, we simulated tensors fields that consist of single and double tensors at the crossing regions. In the first three simulation studies, we examine the dependence of the estimation on the FA and MD of two simulated tensors and angle between the two principal diffusion orientations. The last simulation study is focused on measuring the goodness of fit of the estimated double tensors. A real DTI image from a healthy human

brain in coronal view is presented. The region of pons and corpus callosum is enlarged and the Bayesian models and BN algorithm are performed in the real data. Extra Gaussian noise is added to the DTI image. We estimate the double tensors again and compare the results between raw data and noisy data. The results show that the Bayesian single and double model with neighbouring information as priors can identify the double tensors effectively and are robust. The Bayesian model is more accurate when two fibre orientations are distinct. When the two fibres orientation is similar, i.e. the angle of the principal direction is small, there may be a large variation of FA, MD of the estimated tensors. Especially, the determinant and MD tend to be larger than its true value. The results are optimal when the FA values are large, i.e. the tensors are more isotropic. When the two tensors have similar MD, the Bayesian models also perform better.

In Chapter 4, we consider the positive semi-definite symmetric property of tensors in the tensor field. we define the median non-Euclidean covariance matrix taking its symmetric semi-positive definiteness into account. The estimation techniques include power Euclidean, root Euclidean, log-Euclidean, Riemannian and Procrustes metrics. The tensor estimators are parameterised when necessary to ensure the symmetric semi-positive definiteness property. The weights, inverse Euclidean distances weighting and exponential weighting functions are also considered. The geodesic anisotropy and procrustes anisotropy have been compared with fractional anisotropy index. The weighted non-Euclidean median tensor estimators are compared with Euclidean mean estimators in simulation studies and real data in terms of the determinant, MD, FA and main orientation. The comparison between different estimations of the median tensor is made from a single covariance tensor estimation to the tensor field smoothing and interpolation. The non-Euclidean distances between each two estimated tensors are also compared in the tables. The non-Euclidean median tensor estimators are used to smooth and interpolate the tensor field with weighting functions in simulation studies and real data. We also validate our various non-Euclidean metrics. The root-mean-square deviation of FA, determinant and orientations of tensors are compared in the corpus callosum region. Extra Gaussian noise is added to the diffusion signals. The Euclidean mean and various non-Euclidean median tensors are used to smooth the tensor field. We compare the results with smoothing original raw tensor field. The results show that the Euclidean estimators are undesirable and problematic in the simulation and real data analysis. The non-Euclidean median tensor estimators improve the illustration of DTI images. The log-Euclidean and Riemannian estimators produce similar results. They have advantages over tensor field processing in terms of MD and determinant. These two estimators are more likely to be influenced by outliers. The Procrustes estimator is appropriate in a small variation data set. The power and root Euclidean estimators obtain similar results. The power Euclidean

estimators have an advantage over tensor field processing in terms of FA, but they are not preferable for processing when the size and orientation are of interest. The root Euclidean median estimator is shown to be preferable overall, especially in terms of MD and orientations. It can effectively exclude outliers and clearly show the important structures of the Brain. Weighted median estimators achieve better results than equal weighted estimators. The data shows that the exponential weighting function is preferable especially in terms of the orientation of tensor. The improvement by weights of estimators depends on the different metrics. Overall, the power Euclidean and root Euclidean techniques are recommended among six methods.

Chapter 5 concludes the research and the future work is also suggested.

5.2 Future work

5.2.1 Tensor field segmentation and regularisation

There is a growing need to develop approaches for diffusion tensor processing. Many methods have been proposed to estimate the diffusion tensor, such as least squares methods, and Bayesian frameworks. However, the estimation of diffusion tensors is noise-sensitive due to artifacts and limitations resulting from DTI measurements. Since the diffusion tensor is positive semi-definite, it is more natural to use the non-Euclidean methods for tensor processing. To carry out non-Euclidean statistical analysis of diffusion tensor data could be an extend to my PhD project. The log-Euclidean, Riemannian and power Euclidean metrics could be applied for tensor field processing (e.g. segmentation and regularisation).

5.2.2 Weights modification

The algorithms for weights modification is a possible research area. One would modify the weights in the multi-tensor model, other for weighted non-Euclidean methods. A main concern for the second algorithm is to avoid over-smoothing in an edge region. Additional, relative measurement could be developed to assess the weighting effects.

5.2.3 Smoothing

How to process images with some level of optimality is an important and generally difficult task in many computer vision applications. For diffusion tensor smoothing, the main difficulty lies in the ability to predict how much smoothing needs to applied at various locations. So, it could be interesting to investigate algorithms which would modify the weights in the multi-tensor model which determine a contribution of individual voxels.

5.2.4 Validation

As the complexity of brain tissues and there is no ground truth for the diffusion tensors in the brain, how to validate the result is one of the most difficult questions in DT-MRI. We have evaluated the size, shape and turning angle of the tensor ellipsoid in anatomical basis. How to validate the various procedures in the tract-base analysis could be my future task. To validate the regressions, we could evaluate the locations of the tensors and how they align in the white matter bundles. The non-Euclidean metrics could be used to define the closure the fibre bundles. Segmentation of fibres is another promising field to explore. After defining the each fibre bundle, the central ellipsoid can be found by distance function, or weighted non-Euclidean metric estimators introduced in chapter 4. These ellipsoids represent their fibre bundles. Automatically clustering of tracts concludes fibre bundles by these ellipsoids or FA values. The Euclidean metric and non-Euclidean metrics can be used to compute the distances two fibre bundles. The robustness of the fibres can also be tested. Another validation method is to find the variance of fibre bundles orientations or to define the convergence or divergence covariance tensors of the voxels in each fibre bundles.

5.2.5 Fibre tractography

Another future area of research is to develop a robust statistical analysis on the assessment tool for fibre tractography and to create measures of fibre orientation uncertainty would be of great benefit to neuroscientists. The non-Euclidean metrics we defined in Chapter 4 can be used in white matter tractography. In particular, it is of great interest to consider the connectivity of the brain, and how connectivity analysis is affected by using different metrics and methods. A statistical assessment tool could be developed for fibre tractography. Metric learning methods and new connectivity measures would be developed in the future study.

Bibliography

[1] Alexander, A. L., Lee, J. E., Lazar, M., and Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics*, 4(3):316–329.

[2] Alexander, A. L., Tsuruda, J. S., and Parker, D. L. (1997). Elimination of eddy current artifacts in diffusion-weighted echo-planar images: the use of bipolar gradients. *Magnetic Resonance in Medicine*, 38(6):1016–1021.

[3] Alexander, D. C. (2005). Multiple-fibre reconstruction algorithms for diffusion mri. *Annals of the New York Academy of Sciences*, 1046(1):113–133.

[4] Anscombe, F. J. (1960). Rejection of outliers. *Technometrics*, 2(2):123–146.

[5] Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2005). Fast and simple calculus on tensors in the log-euclidean framework. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 115–122. Springer.

[6] Ashburner, J. (2009). Computational anatomy with the spm software. *Magnetic resonance imaging*, 27(8):1163–1174.

[7] Basser, P. J. and Jones, D. K. (2002). Diffusion-tensor mri: theory, experimental design and data analysis - a technical review. *NMR in Biomedicine*, 15(7-8):456–467.

[8] Basser, P. J., Mattiello, J., and Bihan, D. L. (1994). Mr diffusion tensor spectroscopy and imaging. *Biophysical Journal, Jan.*, 66(1):259–267.

[9] Basser, P. J. and Pajevic, S. (1999). Method to reduce eigenvalue sorting bias in dt-mri. In *Proceedings of the 7th Annual Meeting of ISMRM, Philadelphia, PA*, page 1788.

[10] Batchelor, M., Guan, X.-W., and Oelkers, N. (2004). Thermal and magnetic properties of spin-1 magnetic chain compounds with large single-ion and in-plane anisotropies. *Physical Review B*, 70(18):184408.

[11] Bazin, P.-L., Cuzzocreo, J. L., Yassa, M. A., Gandler, W., McAuliffe, M. J., Bassett, S. S., and Pham, D. L. (2007). Volumetric neuroimage analysis extensions for the mipav software package. *Journal of neuroscience methods*, 165(1):111–121.

[12] Bernardo, J. M. and Smith, A. (2009). Bayesian theory, vol. 405.

[13] Beutel, J., Kundel, H. L., and Metter, R. L. V. (2000). Handbook of medical imaging: Physics and psychophysics. SPIE.

[14] Bigler, D. C., Aksu, Y., Miller, D. J., and Yang, Q. X. (2009). Stamps: Software tool for automated mri post-processing on a supercomputer. *Computer methods and programs in biomedicine*, 95(2):146–157.

[15] Bihan, D. L., Mangin, J. F., Poupon, C., Clark, C. A., S. Pappata, N. M., and Chabriat, H. (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13(4):534–546.

[16] Brang, D., Taich, Z. J., Hillyard, S. A., Grabowecky, M., and Ramachandran, V. S. (2013). Parietal connectivity mediates multisensory facilitation. *Neuroimage*, 78:396–401.

[17] Cárdenes, R., Muñoz-Moreno, E., Tristan-Vega, A., and Martin-Fernandez, M. (2010). Saturn: a software application of tensor utilities for research in neuroimaging. *Computer methods and programs in biomedicine*, 97(3):264–279.

[18] Chandrasekaran, R. and Tamir, A. (1989). Open questions concerning weiszfeld's algorithm for the fermat-weber location problem. *Mathematical Programming*, 44(1-3):293–295.

[19] Chang, L., Jones, D. K., and Pierpaoli, C. (2005). Restore: robust estimation of tensors by outlier rejection. *Magnetic Resonance in Medicine*, 53(5):1088–1095.

[20] Clayden, J. D., Maniega, S. M., Storkey, A. J., King, M. D., Bastin, M. E., Clark, C. A., et al. (2011). Tractor: Magnetic resonance imaging and tractography with r. *Journal of Statistical Software*, 44(8):1–18.

[21] Cohen, M. B., Lee, Y. T., Miller, G., Pachocki, J., and Sidford, A. (2016). Geometric median in nearly linear time. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 9–21. ACM.

[22] Cook, P., Bai, Y., Nedjati-Gilani, S., Seunarine, K., Hall, M., Parker, G., and Alexander, D. (2006). Camino: open-source diffusion-mri reconstruction and processing. In *14th scientific meeting of the international society for magnetic resonance in medicine*, volume 2759. Seattle WA, USA.

[23] Cox, R. W. (1996). Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, 29(3):162–173.

[24] Dempsey, M. F., Condon, B., and Hadley, D. M. (2002). Mri safety review. In *Seminars in Ultrasound, CT and MRI*, volume 23, pages 392–401. Elsevier.

[25] Dryden, I. L., Koloydenko, A., and Zhou, D. (2009). Non-euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Annals of Applied Satatistics*, 3(3):1102–1123.

[26] Dryden, I. L. and Mardia, K. V. (1998). *Statistical shape analysis*, volume 4. J. Wiley Chichester.

[27] Edwards, W., Lindman, H., and Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological review*, 70(3):193.

[28] Fischl, B. (2012). Freesurfer. *Neuroimage*, 62(2):774–781.

[29] Fletcher, P. T. and Joshi, S. (2007). Riemannian geometry for the statistical analysis of diffsion tensor data. *Signal Process*, 87(2):250–262.

[30] Fletcher, P. T., Venkatasubramanian, S., and Joshi, S. (2009). The geometric median on riemannian manifolds with application to robust atlas estimation. *NeuroImage*, 45(1):143–152.

[31] Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de I'institut Henri Poincaré*, 10(4):215–310.

[32] Garyfallidis, E., Brett, M., Amirbekian, B., Nguyen, C., Yeh, F.-C., Halchenko, Y., and Nimmo-Smith, I. (2011). Dipy–a novel software library for diffusion mr and tractography. In *17th annual meeting of the organization for human brain mapping*, pages 1–5.

[33] Goebel, R. and Jansma, H. (2006). Brain voyager qx (version 2.2). *Computer Software and Manual]. Maastricht, The Netherlands: Brain Innovation BV*.

[34] Gupta, R. K., Saksena, S., Agarwal, A., Hasan, K. M., Husain, M., Gupta, V., and Narayana, P. A. (2005). Diffusion tensor imaging in late posttraumatic epilepsy. *Epilepsia*, 46(9):1465–1471.

[35] Guttman, M. A., Zerhouni, E. A., and McVeigh, E. R. (1997). Analysis of cardiac function from mr images. *IEEE Computer Graphics and Applications*, 17(1):30–38.

[36] Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346):383–393.

[37] Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Stahel, W. A. (2011). *Robust statistics: the approach based on influence functions*, volume 114. John Wiley & Sons.

[38] Hogg, R. V. (1979). Statistical robustness: One view of its use in applications today. *The American Statistician*, 33(3):108–115.

[39] Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 221–233.

[40] Huber, P. J. (1972). The 1972 wald lecture robust statistics: A review. *The Annals of Mathematical Statistics*, pages 1041–1067.

[41] Huber, P. J. (1973). Robust regression: asymptotics, conjectures and monte carlo. *The Annals of Statistics*, 1:799–821.

[42] Huber, P. J. (2011). *Robust statistics*. Springer.

[43] Huber, P. J. et al. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1):73–101.

[44] Huber, P. J., Huber, P., Huber, P., Statisticien, M., Suisse, E. U., Huber, P., and Statistician, M. (1996). *Robust statistical procedures*, volume 68. SIAM.

[45] Jaeckel, L. A. (1972). Estimation regression coefficients by minimizing the dispersion of residuals. *The Annual of Mathematical Statistics*, 5:1449–1458.

[46] James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proc. 4th Berkeley Symposium on Mathematical Statistics and Probability*, pages 361–379.

[47] Jara, H. and Wehrli, F. W. (1994). Determination of background gradients with diffusion mr imaging. *Journal of Magnetic Resonance Imaging*, 4(6):787–797.

[48] Jefferys, W. H. and Berger, J. O. (1992). Ockham's razor and bayesian analysis. *American Scientist*, 80(1):64–72.

[49] Jeffrey, H. (1961). Theory of probability. iii edition (i edition 1939, ii edition 1994).

[50] Jeong, Y. Y., Kang, H. K., Chung, T. W., Seo, J. J., and Park, J. G. (2003). Uterine cervical carcinoma after therapy: Ct and mr imaging findings 1. *Radiographics*, 23(4):969–981.

[51] Jeurissen, B., Leemans, A., Tournier, J.-D., Jones, D. K., and Sijbers, J. (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human brain mapping*, 34(11):2747–2766.

[52] Jian, B. and Vemuri, B. C. (2007). A unified computational framework for deconvolution to reconstruct multiple fibers from diffusion weighted mri. *IEEE transactions on medical imaging*, 26(11):1464–1471.

[53] Jiang, H., van Zijl, P. C., Kim, J., Pearlson, G. D., and Mori, S. (2006). Dtistudio: resource program for diffusion tensor computation and fiber bundle tracking. *Computer methods and programs in biomedicine*, 81(2):106–116.

[54] Koay, C. G., Carew, J. D., Alexander, A. L., Basser, P. J., and Meyerand, M. E. (2006a). Investigation of anomalous estimates of tensor-derived quantities in diffusion tensor imaging. *Magnetic Resonance in Medicine*, 55(4):930–936.

[55] Koay, C. G., Chang, L., Carew, J. D., Pierpaoli, C., and Basser, P. J. (2006b). A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging. *Journal of Magnetic Resonance*, 182(1):115–125.

[56] Kuhn, H. W. (1973). A note on fermat's problem. *Mathematical programming*, 4(1):98–107.

[57] Kuo, L., Haldar, J., Lo, Y., Liu, C., Liang, Z., and Tseng, W. (2010). Quantitative improvement of diffusion spectrum imaging tractography using statistical denoising. In *Proceedings of the 18th Annual Meeting of ISMRM, Stockholm, Sweden*, page 1669.

[58] Landman, B. A., Farrell, J. A., Patel, N., Mori, S., and Prince, J. L. (2007). Dti fiber tracking: the importance of adjusting dti gradient tables for motion correction. catnap-a tool to simplify and accelerate dti analysis. In *Proc. org human brain mapping 13th annual meeting*, volume 138.

[59] Lauterbur, P. C. (1973). Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature*, 242(5394):190–191.

[60] Lauterbur, P. C. (1974). Magnetic resonance zeugmatography. *Pure and Applied Chemistry*, 40(1-2):149–157.

[61] Le Bihan, D. (1994). [diffusion, perfusion and functional magnetic resonance imaging.]. *Journal des maladies vasculaires*, 20(3):203–214.

[62] Le Bihan, D. and Iima, M. (2015). Diffusion magnetic resonance imaging: what water tells us about biological tissues. *PLoS Biol*, 13(7):e1002203.

[63] Lee, P. M. (2012). *Bayesian statistics: an introduction*. John Wiley & Sons.

[64] Leemans, A., Jeurissen, B., Sijbers, J., and Jones, D. (2009). Exploredti: a graphical toolbox for processing, analyzing, and visualizing diffusion mr data. In *17th Annual Meeting of Intl Soc Mag Reson Med*, volume 209, page 3537.

[65] Lei, D., Ma, J., Shen, X., Du, X., Shen, G., Liu, W., Yan, X., and Li, G. (2012). Changes in the brain microstructure of children with primary monosymptomatic nocturnal enuresis: a diffusion tensor imaging study. *PLoS One*, 7(2):e31023.

[66] Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4):764–766.

[67] Liu, C.-Y., Iglesias, J. E., Tu, Z., Initiative, A. D. N., et al. (2013). Deformable templates guided discriminative models for robust 3d brain mri segmentation. *Neuroinformatics*, 11(4):447–468.

[68] Loening, A. M., Gambhir, S. S., et al. (2003). Amide: a free software tool for multimodality medical image analysis. *Molecular imaging*, 2(3):131–137.

[69] Lucas, B. C., Bogovic, J. A., Carass, A., Bazin, P.-L., Prince, J. L., Pham, D. L., and Landman, B. A. (2010). The java image science toolkit (jist) for rapid prototyping and publishing of neuroimaging software. *Neuroinformatics*, 8(1):5–17.

[70] MacKay, D. J. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

[71] Maximov, I., Grinberg, F., and Shah, N. J. (2001). Robust tensor estimation in diffusion tensor imaging. *Journal of Magnetic Resonance*, 213(1):136–144.

[72] Melnick, E. L. and Everitt, B. S. (2008). *Encyclopedia of quantitative risk analysis and assessment*, volume 1. John Wiley & Sons.

[73] Miller, J. (1991). Reaction time analysis with outlier exclusion: Bias varies with samplesize. *The Quarterly Journal of Experimental Psychology*, 43(4):907–912.

[74] Mishra, P. B., Bharadwaj, K. G., Chakrabarti, D., Menon, P., Lobo, A. S., Padigaru, M., and Rohatagi, S. (2012). Automated background correction for batch quantification of images using a new imagej macro. *Journal of Computational Intelligence in Bioinformatics*, 5(3):169–174.

[75] Mori, S. and van Zijl, P. (2002). Fiber tracking: principles and strategies–a technical review. *NMR in Biomedicine*, 15(7-8):468–480.

[76] Moseley, M. E., Cohen, Y., Kucharczyk, J., Mintorovitch, J., Asgari, H., Wendland, M., Tsuruda, J., and Norman, D. (1990). Diffusion-weighted mr imaging of anisotropic water diffusion in cat central nervous system. *Radiology*, 176(2):439–445.

[77] Papademetris, X., Jackowski, M. P., Rajeevan, N., DiStasio, M., Okuda, H., Constable, R. T., and Staib, L. H. (2006). Bioimage suite: An integrated medical image analysis suite: An update. *The insight journal*, 2006:209.

[78] Park, H.-J., Kubicki, M., Westin, C.-F., Talos, I.-F., Brun, A., Peiper, S., Kikinis, R., Jolesz, F. A., McCarley, R. W., and Shenton, M. E. (2004). Method for combining information from white matter fiber tracking and gray matter parcellation. *American Journal of Neuroradiology*, 25(8):1318–1324.

[79] Paty, D., Oger, J., Kastrukoff, L., Hashimoto, S., Hooge, J., Eisen, A., Eisen, K., Purves, S., Low, M., Brandejs, V., et al. (1988). Mri in the diagnosis of ms a prospective study with comparison of clinical evaluation, evoked potentials, oligoclonal banding, and ct. *Neurology*, 38(2):180–180.

[Pennec et al.] Pennec, X., Fillard, P., and Ayache, N. A riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66.

[81] Pieper, S., Halle, M., and Kikinis, R. (2004). 3d slicer. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 632–635. IEEE.

[82] Pierpaoli, C. and Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic resonance in Medicine*, 36(6):893–906.

[83] Pierpaoli, C., Walker, L., Irfanoglu, M., Barnett, A., Basser, P., Chang, L., Koay, C., Pajevic, S., Rohde, G., Sarlls, J., et al. (2010). Tortoise: an integrated software package for processing of diffusion mri data. *Book TORTOISE: an Integrated Software Package for Processing of Diffusion MRI Data (Editor edˆ eds)*, 18:1597.

[84] Pinheiro, J. C. and Bates, D. G. (1996). Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing*, 6(3):289–296.

[85] Purcell, E. M., Torrey, H. C., and Pound, R. V. (1946). Resonance absorption by nuclear magnetic moments in a solid. *Physical review*, 69(1-2):37.

[86] Riani, M. and Atkinson, A. C. (2000). Robust diagnostic data analysis: Transformations in regression. *Technometrics*, 42(4):384–494.

[87] Riani, M. and Atkinson, A. C. (2006). Distribution theory and simulations for tests of outliers in regression. *Journal of Computational and Graphical Statistics*, 15(2):460–476.

[88] Riani, M., Atkinson, A. C., and Cerioli, A. (2009). Finding an unknown number of multivariate outliers. *Journal of the Royal Statistical Society: series B (statistical methodology)*, 71(2):447–466.

[89] Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of American Statistical Association*, 79(388):871–880.

[90] Rousseeuw, P. J. (1991). Tutorial to robust statistics. *Journal of Chemometrics*, 5(1):1–20.

[91] Rousseeuw, P. J. and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283.

[92] Rousseeuw, P. J. and Hubert, M. (1997). Recent developments in progress. *Lecture Notes–Monograph Series*, 31:201–214.

[93] Rousseeuw, P. J. and Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. John Wiley & Sons, ISBN: 0471852333.

[94] Schellinger, P. D., Thomalla, G., Fiehler, J., Köhrmann, M., Molina, C. A., Neumann-Haefelin, T., Ribo, M., Singer, O. C., Zaro-Weber, O., and Sobesky, J. (2007). Mri-based and ct-based thrombolytic therapy in acute stroke within and beyond established time windows. *Stroke*, 38(10):2640–2645.

[95] Shen, D. and Davatzikos, C. (2002). Hammer: hierarchical attribute matching mechanism for elastic registration. *IEEE transactions on medical imaging*, 21(11):1421–1439.

[96] Sherbondy, A., Akers, D., Mackenzie, R., Dougherty, R., and Wandell, B. (2005). Exploring connectivity of the brain's white matter with dynamic queries. *IEEE transactions on visualization and computer graphics*, 11(4):419–430.

[97] Siegel, A. F. (1982). Robust regression using repeated medians. *Biometrika*, 69(1):242–244.

[98] Small, G. and Vorgan, G. (2009). ibrain: Surviving the technological alteration of the modern mind. *Education Review//Reseñas Educativas*.

[99] Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage*, 31(4):1487–1505.

[100] Souplet, J.-C., Fillard, P., Toussaint, N., Frenay, C. L., and Malandain, G. (2008). Sepinria: a software to analyse multiple sclerosis brain mri. In *Neurology*, volume 70, pages A474–A474. LIPPINCOTT WILLIAMS & WILKINS TWO COMMERCE SQ, 2001 MARKET ST, PHILADELPHIA, PA 19103 USA.

[101] Stachler, R. J., Chandrasekhar, S. S., Archer, S. M., Rosenfeld, R. M., Schwartz, S. R., Barrs, D. M., Brown, S. R., Fife, T. D., Ford, P., Ganiats, T. G., et al. (2012). Clinical practice guideline: sudden hearing loss. *OtolaryngologyHead and Neck Surgery*, 146(3_suppl):S1–S35.

[102] Stigler, S. M. (1981). Gauss and the invention of least squares. *The Annals of Statistics*, pages 465–474.

[103] Torti, F., Perrotta, D., Atkinson, A. C., and Riani, M. (2012). Benchmark testing of algorithms for very robust regression: Fs, lms and lts. *Computational Statistics & Data Analysis*, 56(8):2501–2512.

[104] Toussaint, N., Souplet, J.-C., Fillard, P., et al. (2007). Medinria: medical image navigation and research tool by inria. In *Proc. of MICCAI*, volume 7, page 280. Citeseer.

[105] Tuch, D. S. (2004). Q-ball imaging. *Magnetic resonance in medicine*, 52(6):1358–1372.

[106] Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., and Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48(4):577–582.

[107] Vardi, Y. and Zhang, C.-H. (2000). The multivariate l1-median and associated data depth. *Proceedings of the National Academy of Sciences*, 97(4):1423–1426.

[108] Wang, R., Benner, T., Sorensen, A., and Wedeen, V. (2007). Diffusion toolkit: a software package for diffusion imaging data processing and tractography. In *Proc Intl Soc Mag Reson Med*, volume 15.

[109] Wang, Y., Gupta, A., Liu, Z., Zhang, H., Escolar, M. L., Gilmore, J. H., Gouttard, S., Fillard, P., Maltbie, E., Gerig, G., et al. (2011). Dti registration in atlas based fiber analysis of infantile krabbe disease. *Neuroimage*, 55(4):1577–1586.

[110] Wedeen, V., Reese, T., Tuch, D., Weigel, M., Dou, J., Weiskoff, R., and Chessler, D. (2000). Mapping fiber orientation spectra in cerebral white matter with fourier-transform diffusion mri. In *Proceedings of the 8th Annual Meeting of ISMRM, Denver*, page 82.

[111] Wu, X.-L., Gai, J., Lam, F., Fu, M., Haldar, J. P., Zhuo, Y., Liang, Z.-P., Hwu, W.-M., and Sutton, B. P. (2011). Impatient mri: Illinois massively parallel acceleration toolkit for image reconstruction with enhanced throughput in mri. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 69–72. IEEE.

[112] ZHANG, T.-l., LI, C.-y., LIU, H.-j., et al. (2010). Measurement of the volume of corpus callosum and lateral ventricle in normal adults and in patients with vascular dementia. *Radiologic Practice*, 12:017.

[113] Zhou, D. (2010). *Statistical analysis of diffusion tensor imaging*. PhD thesis, University of Nottingham.